```
09 327, 984
```

GenCore version 4.5 Copyright (c) 1993 - 2000 Compugen Ltd.

OM protein - protein search, using sw model

Run on: January 15, 2002, 13:50:09; Search time 120.01 Seconds (without alignments) 357.991 Million cell updates/sec

Title: US-09-724-296-38 Perfect score: 3055

Sequence: 1 MGTTGLESLSLGDRGAAPTV......HKLNKYHDWLWENHEKLSLS (580)
Scoring table: BLOSUM62

Scoring table: BLOSUM62 Gapop 10.0 , Gapext 0.5 Searched: 522463 segs, 74073290 residues Total number of hits satisfying chosen parameters: 522463

Minimum DB seq length: 0 Maximum DB seq length: 200000000 Post-processing: Minimum Match 0% Maximum Match 100% Listing first 45 summaries

Database : A_Gen

A_Geneseq_1101:* .: /SIDSB/gcgdata/geneseq/geneseqp/AA1980.DAT:* /SIDS8/gcgdata/geneseg/genesegp/AA2000.DAT: /SIDS8/gcgdata/geneseg/genesegp/AA2001.DAT: /SIDS8/gcgdata/geneseq/geneseqp/AA1988.DAT: /SIDS8/gcgdata/geneseq/geneseqp/AA1994. /SIDS8/qcgdata/geneseg/geneseqp/AA1995. /SIDS8/gcgdata/geneseg/genesegp/AA1998. /SIDS8/gcgdata/geneseg/genesegp/AA1999. /SIDS8/gcgdata/geneseq/geneseqp/AA1989 /SIDS8/gcgdata/geneseq/genesegp/AA1990 /SIDS8/gcgdata/geneseg/genesegp/AA1991. /SIDS8/gcgdata/geneseg/genesegp/AA1992. /SIDS8/gcgdata/geneseq/geneseqp/AA1985.1 /SIDS8/gcgdata/geneseq/geneseqp/AA1986.1 /SIDS8/gcgdata/geneseqp/AA1987.1 /gcgdata/geneseg/genesegp/AA1995 /SIDS8/gcgdata/geneseq/geneseqp/AA1996 /SIDS8/gcgdata/geneseq/geneseqp/AA1997 /SIDS8/gcgdata/geneseq/geneseqp/AA1993 /SIDS8/gcgdata/geneseq/geneseqp/AA1983. /SIDS8/gcgdata/geneseq/geneseqp/AA1984. /SIDS8/gcgdata/geneseq/geneseqp/AA1982. /SIDS8/gcgdata/geneseq/geneseqp/AA1981

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Description	Human MED1 endonuc	Human delta228-UV	Human ovarian tumo	Secreted protein e	Human secreted pro	Arabidopsis thalia	Arabidopsis thalia	Arabidopsis thalia	BRCA2 cancer susce	Human breast cance	Human breast and o
QI S	AAW74473	AAY44504	AAY76548	AAW88701	AAG02051	AAG16859	AAG16858	AAG16857	AAW25049	. AAW19211	AAW23287
80	20	21	20	20	21	7	21	71	18	18	18
% Query Match Length DB ID	580	565	439	257	219	326	327	352	1644	3418	3418
% Query Match	100.0	8.96	71.8	40.3	6.4	4.9	4.9	4.9	4.7	4.7	4.7
Score	3055	2958.5	2194.5	1230	194	148.5	148.5	148.5	145	145	145
Result No.	-	7	m	4	S	ω	7	ω,	6	10	11

New nucleic acid encoding human endonuclease MED1 involved in DNA mismatch repair - used for diagnosing susceptibility to cancer and

(FOXC-) FOX CHASE CANCER CENT.

WPI; 1999-142462/12. N-PSDB; AAX22002#.

Bellacosa A;

,	AAYO4356 Human BRCA2 (Omi3)	Himan BRCAZ	Human BRCA2	BRCA2 proteir	Par	Hum	10.1		Hum	Tat	HUH HUH		Hea	dxa	E	H. pylori GHPO	Human prostate	Human prostat	Mouse DNA der		Peptide #4097	Arabidopsis	Arabidopsis	Arabidopsis	Canine ract	ממק שחוולשטי	Peptide #941	Human ORFX	Plasmodium f	ALIGNMENTS			580 AA.				protein sequence.	methyl-Cod hinding endomiolesce-1.	cancer; f	disease; spinocerebellar ataxia;	0	6	1 P. P. J. 1800			mark mark .		
	18 20	10	7	7		~ ~	10		7		71 (4 (2	7	П,	٦ ،	A (د	1 (7)	7	~	2	(4)	~ (4 C	4 (1 (1	7	7				Protein;		1	encry)	αu	human. m	ipulati	untingt	ipiet					8WO-US15828.	00000	/US-UU3336.
	.7 34	7 34	.7 34	.7 34	.7 23	.7 34	. L	. e.	.3	w.	J. C.	. m.	· Ψ.	.3	e	٤٠ د ر	. c	4 C	. 0.	0.	.0 12	0.	0.0	0.	v.	90	.9	.9	6.				andard; Pr			a lirst e	endonucleas			cophy; H	ase; tr					98WO-U	0 2000	0-50/6
,	2 145 4	145	145	145	144	7,7	137	131.5	131.5	131	130.5	130.5	130.5	130	130	130	27.75	126	123	122.5	122.5	121.5	121.5	121.5	120.5	120.5	120	120.5	119	House		r 1 473	AAW74473 stand	AAW74473;	000) 8661-1AM-61	Human MED1 endonucle	י שפפס נטוומטטמפ	ONA fidelity;	nyotonic dystr	rennedy's dist	Homo sapiens.	W09904626-A1		04-FEB-1999.	28-JUL-1998;	, 1002 - HIT	; /66T-700-87
,	12	4.	15	16	17	B 7 F	20	21	22	23	400	26	27	28	. 500) L	7 6	י מ	34	35	36	, m	m c	, c	. 7	• •	4	44	4	Mary in Commerce and Assessment	A COLUMN TO A COLU	RESULT AAW744		AX AC		X						SOS		(-	11	XX	

			** · · · · · · · · · · · · · · · · · ·

```
for DNA manipulation and to reason antibodies. Susceptibility or predisposition to cancer (particularly colorectal or endometrial, especially heredistry on replyposis colorectal cancer, or its prognosis, where caused by alterations in the MEDI-encoding gene, are identified by sequence comparison, amplification, detecting altered polypeptide, and restriction fragment mapping, hybridisation (particularly to probes specific for a mutant allele). These same methods can also be used to diagnose fragile X syndrome and other diseases (e.g., myotonic dystrophy, Huntington's disease, spinocerebellar ataxia and Kennedy's disease) associated with triplet repeat expansion. The DNA, or its fragments, are used as probes and primers in the above diagnostic methods, also to isolate homologous sequences, as sources of antisense sequences and for
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    gene transfer, particularly to restore drug sensitivity to drug-resistant
                                                                                                                                                                           MED1.(for methyl-CpG binding endonuclease-1) is used to screen for specific modulators (potential therapeutic agents particularly mimetics of MED1) and to study interactions involved in maintaining DNA fidelity,
                                                                                                                                                   This sequence is the human MED1 endonuclease of the invention.
                                 fragile X syndrome, and therapeutically
                                                                                          Claim 8; Fig 3; 109pp; English.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               cancer cells.
*£$8$66666666666666688888
```

Human delta228-UV damage endonuclease.

27-MAR-2000 (first entry)

AAY44504;

AAY44504 standard; Protein; 565

580 AA; Sequence

ö 121 KSSLANYLHKNGETSLKPEDFDFTVLSKRGIKSRYKDCSMAALTSHLQNQSNNSNWNLRT 180 RSKCKKDVFMPPSSSSELQESRGLSNFTSTHLLLKEDEGVDDVNFRKVRKPKGKVTILKG 240 241 IPIKKTKKGCRKSCSGFVQSDSKRESVCNKADAESEPVAQKSQLDRTVCISDAGACGETL 300 241 ipikktkkgcrkscsgfvgsdskresvcnkadaesepvaqksgldrtvcisdagacgetl 300 SVTSEENSLVKKKERSLSSGSNFCSEQKTSGIINKFCSAKDSEHNEKYEDTFLESEEIGT 360 KVEVVERKEHLHTDILKRGSEMDNNCSPTRKDFTGEKIFQEDTIPRTQIERRKTSLYFSS 420 KYNKEALSPPRRKAFKKWTPPRSPFNLVQETLFHDPWKLLIATIFLNRTSGKMAIPVLWK 480 481 FLEKYPSAEVARTADWRDVSELLKPLGLYDLRAKTIVKFSDEYLTKQWKYPIELHGIGKY 540 61 CNPLLQEPIASAQFGATAGTECRKSVPCGWERVVKQRLFGKTAGRFDVYFISPQGLKFRS 120 Gaps 1 MGTTGLESLSLGDRGAAPTVTSSERLVPDPPNDLRKEDVAMELERVGEDEEQMMIKRSSE 60 1 mgttgles1s1gdrgaaptvtsserlvpdppndlrkedvamelervgedeegmmikrsse 60 ; 0 100.0%; Score 3055; DB 20; Length 580; 100.0%; Pred. No. 7.6e-284; Indels 541 GNDSYRIFCVNEWKQVHPEDHKLNKYHDWLWENHEKLSLS 580 Mismatches ő Conservative Similarity Matches 580; Query Match Best Local 421 121 301 361 181 61 g g ద a a ò g ò a. ŏ a ŏ g õ ò ò ö

```
The present sequence is human delta228-UV damage endonuclease.
Delta228-UVDE is a 288 amino acid deletion of the N-terminal of the
uvel+ gene product. This is expressed in frame with a GST leader sequence
to generate a fusion protein. This provides stable endonuclease fragments
for cleaving a double-stranded DNA molecule that has a distorted
structure resulting from UV radiation damage, a photoproduct, an abasic
site, mismatched nucleotide pairing, aplatinum diaduct, an intercalated
molecule or alkylation of a nucleotide. Uvelp can be used in compositions
for internal or topical application and as a therapeutic agent for skin
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    new truncated ultraviolet damage endonuclease for treatment of skin
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   61 CNPLLQEPIASAQFGATAGTECRKSVPCGWERVVKQRLFGKTAGRFDVYFISPQGLKFRS 120
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 241 IPIKKTKKGCRKSCSGFVQSDSKRESVCNKADAESEPVAQKSQLDRTVCISDAGACGETL 300
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   Gaps
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                1 MGTTGLESLSLGDRGAAPTVTSSERLVPDPPNDLRKEDVAMELERVGEDEEQMMIKRSSE 60
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 1 mgttgles1s1gdrgaaptvtsser1vpdppndlrkedvamelervgedeegmm1krsse 60
                                                                                                                                          Delta228-UVDE; ultraviolet damage endonuclease; GST signal peptide;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              121 KSSLANYLHKNGETSLKPEDFDFTVLSKRGIKSRYKDCSMAALTSHLQNQSNNSNWNLRT
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            181 RSKCKKDVFMPPSSSSELQESRGLSNFTSTHLLLKEDEGVDDVNFRKVRKPKGKVTILKG
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             Query Match

96.8%; Score 2958.5; DB 21; Length 565;
Best Local Similarity 97.4%; Pred. No. 1.3e-274;
Matches 565; Conservative 0; Mismatches 0; Indels 15;
                                                                                                                                                        glutathione-S-transferase signal peptide; uvel+ gene product; UV irradiation; DNA damage; UV radiation damage; photoproduct; abasic site; aplatinum diaduct; mismatched nucleotide pairing; nucleotide alkylation; skin cancer.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                Claim 16; Page 60; 133pp; English.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                          Avery AM;
                                                                                                                                                                                                                                                                                                                                                                                98US-0088521.
99US-0134752.
                                                                                                                                                                                                                                                                                                                                                   99WO-US12910
                                                                                                                                                                                                                                                                                                                                                                                                                                                                          Kaur B,
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            WPI; 2000-116417/10
                                                                                                                                                                                                                                                                                                                                                                                                                                       (UYEM-) UNIV EMORY.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            565 AA;
                                                                                                                                                                                                                                                                              W09963828-A1
                                                                                                                                                                                                                                                                                                                                                  08-JUN-1999;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                          Doetsch PW,
                                                                                                                                                                                                                                                                                                                                                                                  38-JUN-1998;
                                                                                                                                                                                                                                                                                                                                                                                                    18-MAY-1999;
                                                                                                                                                                                                                                                                                                                 16-DEC-1999,
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             Sequence
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               cancers
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  q
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       qq
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            В
                                                                                                                                                                                                                                                                                                                                                                                    õ
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       δ
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 ð
```

ï

		·	
			4 p

us-09-724-296-2_copy_230_828.rspt

```
.
           ó
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         "A eukaryotic gene encoding an endonuclease that specifically repairs DNA damaged by ultraviolet light."; EMBO J. 14:2393-2399(1995)
                                                                                                                                                                                                                                                                                                                                                                                                     481 IPTIRETWTRKGITQKQHYSESADPTAISGMKRRAHSDRVFDFPPCDPTMDLMIEAKEKE 540
                                                                                                                                                                                                                                                                                                                                                                                                                                                     300
                                                                                                                                                                                                                                                                                                                                                   QRLSDSVKARLVLENDDVSWSVQDLLPLCQELNIPLVLDWHHHNIVPGTLREGSLDLMPL 480
                                                                                                                                                                                                   IASPREVVVDSAIRDLAYHDEILSRMKLNEQLNKDAVLIIHLGGTFEGKKETLDRFRKNY 420
                                                                                      61 TLSMLPQVNIGANSFSAETPVDLKKENETELANISGPHKKSTSTSTRKRARSSKKKATDS 120
                                                                                                                                                                                                                                                                                      EWNHNFGIHFMRVSSDLFPFASHAKYGYTLEFAQSHLEEVGKLANKYNHRLTMHPGQYTQ 360
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   541 QAVFELCRRYELQNPPCPLEIMGPEYDQTRDGYYPPGAEKRLTARKRRSRKEEVEEDEK 599
                                    MLRLLKRNIQISKRIVFTILKQKAFKGNHPCVPSVCTITYSRFHCLPDTLKSLLPMSSKT 60
                                                  1 MLRLLKRNIQISKRIVFTILKQKAFKGNHPCVPSVCTITYSRFHCLPDTLKSLLPMSSKT 60
              Gaps
                                                                                                 VSDKIDESVASYDSSTHLRRSSRSKKPVNYNSSSESESEEQISKATKKVKQKEEEETVEE
                                                                                                                                                   VDEKSLKNESSSDEFEPVVPEQLETPISKRRRSRSSAKNLEKESTMNLDDHAPREMFDCL
                                                                                                                                                                                                                                       DKPIPWRGRLGYACLNTILRSMKERVFCSRTCRITTIQRDGLESVKQLGTQNVLDLIKLV
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        541 QAVFELCRRYELQNPPCPLEIMGPEYDQTRDGYYPPGAEKRLTARKRRSRKEEVEEDEK
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                rajima H., Takao M., Yasuhira S., Zhao J.H., Ishii C., Inoue H.,
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  Neurospora crassa.
Eukaryota; Hungi, Ascomycota; Pezizomycotina; Sordariomycetes;
Sordariales; Sordariaceae; Neurospora.
NCBI_TaxID=5141;
             0
              Indels
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         4B231E7A8735C86F CRC64;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     Created)
Last sequence update)
Last annotation update)
 Pred. No. 7.9e-190;
               .
0
              Mismatches
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               ¥.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          STRAIN=74-OR23-1A;
MEDLINE=95292980; PubMed=7774597;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         656 AA; 74413 MW;
              0
100.08;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               10,
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     01,
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                EMBL; D11392; CAB20113.1;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               PRELIMINARY;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    01-NOV-1996 (TrEMBLrel.
01-MAY-1999 (TrEMBLrel.
01-MAY-1999 (TrEMBLrel.
               Conservative
Best Local Similarity
Matches 599; Conserv
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              SEQUENCE FROM N.A.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           UV - ENDONUCLEASE
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              Endonuclease
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 Rasui A.;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         SEQUENCE
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            Q01408
Q01408;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      7
                                                                                                                                                                                                                                                                241
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       UVE1
                                                                                                                61
                                                                                                                                       121
                                                                                                                                                               121
                                                                                                                                                                                      181
                                                                                                                                                                                                               181
                                                                                                                                                                                                                                        241
                                                                                                                                                                                                                                                                                                                301
                                                                                                                                                                                                                                                                                                                                         361
                                                                                                                                                                                                                                                                                                                                                                  361
                                                                                                                                                                                                                                                                                                                                                                                         421
                                                                                                                                                                                                                                                                                        301
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      RESULT
Q01408
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                d
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                g
                                                                                                                                                                                                                                                             a
                                                                                                                                                                                                                                                                                                             g
                                                                                                                                                                                                                                                                                                                                         ò
                                                                                                                                                                                                                                                                                                                                                               g
                                                                                                                                                                                                                                                                                                                                                                                        ò
                                                                                                                                                                                                                                                                                                                                                                                                                                           à
                                                                                                               g
                                                                                                                                                                                                             g
                                                                                                                                                                                                                                       ò
                                                                                                                                                                                                                                                                                        ò
                                                                                       ŏ
                                                                                                                                       ò
                                                                                                                                                            g
                                                                                                                                                                                      δ
```

Length 656;

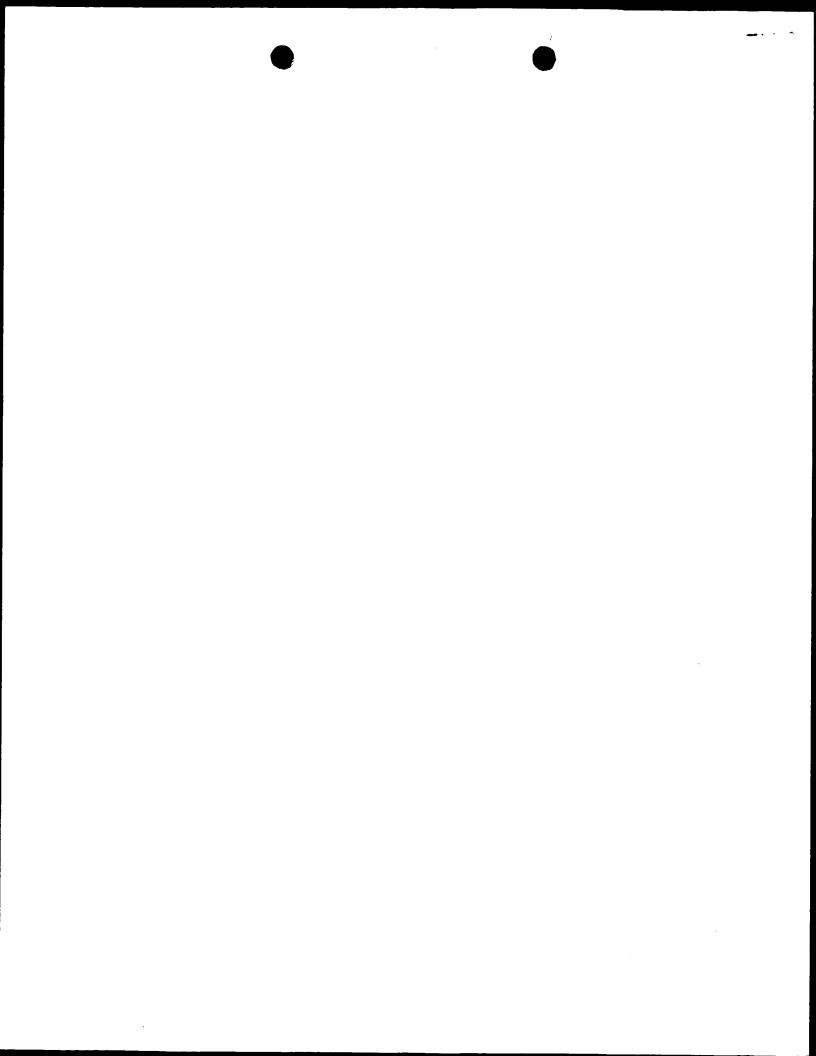
Score 1011.5; DB 3; Pred. No. 2.2e-56;

32.5%; 36.6%;

Local Similarity

Query Match Best Local S

```
18;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          Okstad O., Gominet M., Purnelle B., Rose M., Lereclus D., Kolsto A.B.; "Sequence analysis of three Bacillus cereus loci under Pick-regulated genes encoding degradative enzymes and enterotoxin."; Microbiology 145:3129-3138(1999).
EMBL: AJ243712; CAB69813.1; -...
Endonuclease.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          234 SIVDHRHPLQFEDEPEHHLKNKPDKSKEPQDELGHKFVQELGLANARDIVKMLCWNEKYG 293
                                                                                                                                                                                                                                                                                                                           IHFMRVSSDLFPFASHAKYGYTL-EFAQSHLEEVGKLANKYNHRLTMHPGQYTQIASPRE 366
                                                                                                                                                                                                                                                                                                                                            367 VVVDSAIRDLAYHDEILSRMKLNEQLNKDAVLIIHLGGTFEGKKETLDRFRKNYQRLSDS 426
                                                                                                                                                                                                                                                                                                                                                                                                      427 VKARLVLENDDVSWSVQDLLPLCQELNIPLVLDWHHHNIV--PGTLREGSLDLM--PLIP 482
                                                                                                                                                                                                                                                                                                                                                                                                                                                                483 TIRETWTRKGITQKQHYSESADPTAISGMKRRAHSDRVFDFPPCDPTMDLMIEAKEKEQA 542
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                533 VFELMRTFKLPGFEKINDMVPYDRDDENRPAPPVKAPKKKKGGKRKRTTDEEAAEPEEVD 592
                                                                                                       | :: :|||: |:: ||:
AGRMMGKDANGHCLREGKEQEEGVKMAIEGLARMERRLQRATKRQKKQLEED----- 114
                                                                                                                                                                                                                                                                 ------QRDGLESVKQLGTQNVLDLIKLVEWNHNFG 307
                            ARSSKKKATDSVSDKIDESVASYDSST-----HLRRSSR-----SKKPVNYNSSSESE- 157
                                                                                     --SEEQISKATKKVKQKEEEEYVEEVDEK 184
Indels 181; Gaps
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            --YDQTRDG-----EEVEED 597
                                                                                                                                                                                                         ----PREMFDCLDK-------PIPWRGRLGYACLNTILRSMKERVFCSRTCRIT
                                                                                                                                               185 SLKNESSSDEFEPVVPEQLETPISKRRS-----RSSAKNLEKESTMNLDDHA-----
                                                        SRKSKAAALDTPQSESSTFSSTLDSSAPSPARNLRRSGRNILOPSSEKDRDHEKRSGEEL
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              317 AA; 36990 MW; A6AC35F5800A22C3 CRC64;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 Created)
Last sequence update)
Last annotation update)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    Bacteria; Firmicutes; Bacillus/Clostridium group;
Bacillus/Staphylococcus group; Bacillus.
NCBI_TaxID=1396;
88; Mismatches 149;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      -------
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     AA
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     317
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 SEQUENCE FROM N.A.
STRAIN-ATCC 14579 TYPE STRAIN;
MEDLINE-20055637; PubMed-10589720;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    PRT;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                01-OCT-2000 (TrEMBLrel. 15, 01-OCT-2000 (TrEMBLrel. 15, 01-MAR-2001 (TrEMBLrel. 16,
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           PUTATIVE UV-ENDONUCLEASE
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     PRELIMINARY;
   Conservative
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       543 VFELCRRYEL-
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           Bacillus cereus
                                                                                                                                                                                                                                                                     TI
   Matches 241;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           SEQUENCE
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     09L4F0
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     Q9L4F0
                                                                                                                                                                                                           233
                                                                                                                                                                                                                                                                      276
                                                                                                                                                                                                                                                                                                                              308
                                                           m
                                                                                                                      63
                                                                                         158
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       Q9L4F0
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   q
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   à
                                                        pp
                                                                                         δy
                                                                                                                     g
                                                                                                                                                 ò
                                                                                                                                                                             QQ
                                                                                                                                                                                                            ά
                                                                                                                                                                                                                                     q
                                                                                                                                                                                                                                                                     ò
                                                                                                                                                                                                                                                                                                 qq
                                                                                                                                                                                                                                                                                                                                δ
                                                                                                                                                                                                                                                                                                                                                          g
                                                                                                                                                                                                                                                                                                                                                                                          ò
                                                                                                                                                                                                                                                                                                                                                                                                                      g
                                                                                                                                                                                                                                                                                                                                                                                                                                                    ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               δ
```



```
January 15, 2002, 13:52:23; Search time 74.94 Seconds (without alignments) 666.807 Million cell updates/sec
GenCore version 4.5
Copyright (c) 1993 - 2000 Compugen Ltd.
                                                                                     OM protein - protein search, using sw model
                                                                                                                                    Run on:
```

3482 1 MPSRKSKAAALDTPQSESST......REVKKGKVPEEVEDEGEFDG 656 219241 seqs, 76174552 residues BLOSUM62 Gapop 10.0 , Gapext 0.5 US-09-724-296-36 Title: Perfect score: Sequence: Scoring table: Searched:

al number of hits satisfying chosen parameters: Maximum DB seq length: 0
Maximum DB seq length: 2000000000

219241

Post-processing: Minimum Match 0% Maximum Match 100% Listing first 45 summaries Database :

PIR_68:*
1: pir1:*
2: pir2:*
3: pir3:*
4: pir4:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

Description	UV-endonuclease - UV-endonuclease - hypothetical profe	ease h damag l prot	hypothetical prote NF-180 - sea lampr hypothetical prote hypothetical prote IgA-specific metal	blackjack protein, proliferation pote cell proliferation cell proliferation neurofilament trip neurofilament trip bassoon protein hypothetical prote caldesmon human centromere protein hypothetical prote (ESI) protein (NGC-89 - C hypothetical prote neurofilament trip hypothetical protein hypothetical	
SUMMARIES	i • •				
SUMM ID	S55262 S71134 E83974	S55418 C75350 T00365 T13564	151116 117292 133100 C81169	T28657 B48727 B48666 A48666 S0030 T42730 T124866 JH0628 S28261 E71606 E71606 T229757 T229757	
DB	777	00000	10000		
Length	656 599 322	320 326 1280 5327 796	1110 727 1510 1815	1547 12560 12560 12560 12560 12663 12663 12563 1222 1222 1222 1222 1222	
å Query Match			44444		
Score	3472 1011.5 377	348.5 261.5 160.5 155.5	141.5 140 139.5 139	138 137 136.5 136.5 136.5 135.5 133.5 132.5 132.5 132.5 131.5	/
Result No.	H 0 M	4 70 0 7 8	10 11 12 13	22 2 2 2 2 2 2 2 2 2 2 2 2 2 2 2 2 2 2	

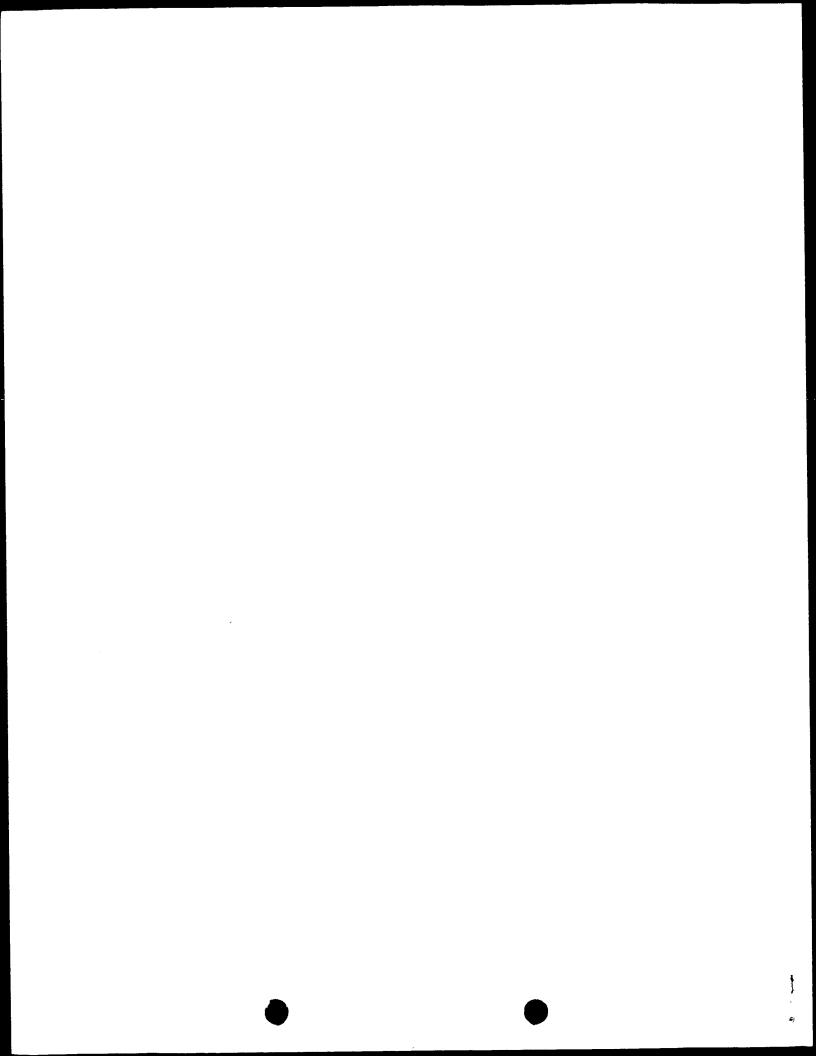
plastic titin . bu	neurofilament trin		hypothetical prote	oronin gen aroto	proofil gene proce	hypothetical proto	adenomatous nolumo	hypothetical proto	Tal-specific motal	hypothetical aroto	hypothetical prote	erythrocute membra	hypothetical proto	aphirin 3 long on	neurofilament trip
I38346	ОЕНОН	A34736	T34418	T13734	560254	126819	T13825	C96608	A81937	C86406	T52359	D71623	T26135	A55575	A37221
~	П	7	~	7	7	7	7	7	7	~	7	7	~	7	7
7962	1020	1805	3488	4151	2453	720	2416	522	1773	786	1400	2441	1829	4377	1072
3.7	3.7	3.7	3.7	3.7	3.7	3.6	3.6	3.6	3.6	3.6	3.6	3.6	3.6	3.6	3.6
128.5 3.7	128 3.7	128 3.7	. 128 3.7	128 3.7	127.5 3.7	126.5 3.6	126.5 3.6	126 3.6	126 3.6	125.5 3.6	125.5 3.6	125.5 3.6	124.5 3.6	124.5 3.6	124 3.6

ALIGNMENTS

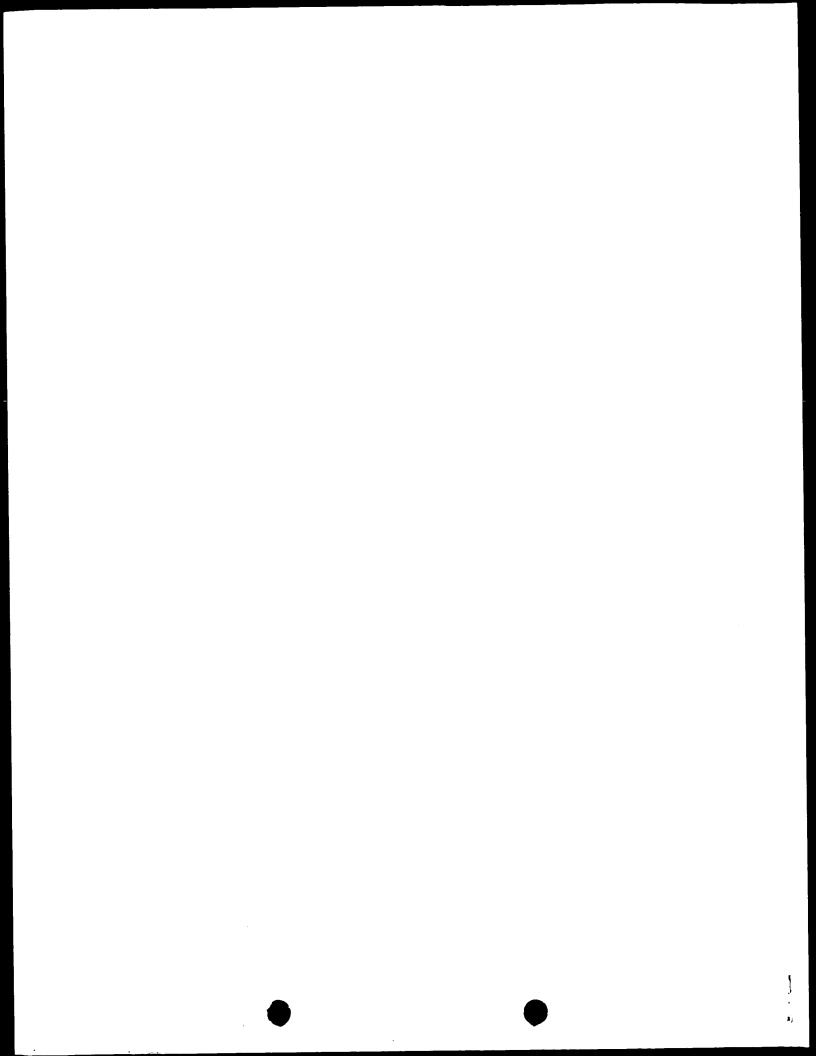
RESULT 1 S55262 UV endonuclease - Neurospora crassa C; Decies: Neurospora crassa C; Dete: 10-Oct-1995 #sequence_revision 03-Nov-1995 #text_change 20-Jun-2000 C; Accession: 555262 R; Yajiana, H; Takao, M; Yasuhira, S.; Zhao, J.H.; Ishii, C.; Inoue, H.; Yasui, A. R; Yajiana, H; Takao, M; Yasuhira, S.; Zhao, J.H.; Ishii, C.; Inoue, H.; Yasui, A. R; Yajiana, H.; Takao, M; Yasuhira, S.; Zhao, J.H.; Ishii, C.; Inoue, H.; Yasui, A. R; Yajiana, H.; Takao, M; Yasuhira, S.; Zhao, J.H.; Ishii, C.; Inoue, H.; Yasui, A. R; Yajiana, H.; Takao, M; Yasuhira, S.; Zhao, J.H.; Ishii, C.; Inoue, H.; Yasui, A. R; Yajiana, H.; Takao, M; Yasuhira, S.; Zhao, J.H.; Ishii, C.; Inoue, H.; Yasui, A. R; Yajiana, H.; Takao, M; Yasuhira, S.; Zhao, J.H.; Ishii, C.; Inoue, H.; Yasui, A. R; Yajiana, H.; Takao, M; Yasuhira, S.; Zhao, J.H.; Ishii, C.; Inoue, H.; Yasui, A. R; Yajiana, H.; Yasui, A. R; Yasiana, H.; Yasui, A. R; Yajiana, H.; Yasui, A. R; Yasui, A. R	Query Match Best Local Similarity 99.7%; Pred. No. 2.2e-208; Matches 654; Conservative 1; Mismatches 1; Indels 0; Gaps 0;	1 MPSRKSKAAALDTPQSESSTFSSTLDSSAPSPARNLRRSGRNILQPSSEKDRDHEKRSGE 60 	61 ELAGRNMGKDANGHCLREGKEQEEGVKMAIEGLARMERRLQRATKROKKQLEEDGIPVPS 120 	121 VVSRPPTAPYHHKSTNAEEREAKEPVLKTHSKDVEREAEIGVDDVVKMEPAATNIIEPED 180 	181 AQDAAERGAARPPAVNSSYLPLPWKGRLGYACLNTYLRNAKPPIFSSRTCRMASIVDHRH 240 	241 PLQFEDEPEHHLKNKPDKSKEPQDELGHKFVQELGLANARDIVKMLCWNEKYGIRFLRLS 300 	301 SEMEPFASHPVHGYKLAPFASEVLAEAGRVAAELGHRLTTHPGOFTQLGSPRKEVVESAI 360 	361 RDLEYHDELLSLLKLPEQONRDAVMIIHMGGOFGDKAATLERFKRNYARLSQSCKNRLVL 420
RESULT 1 S55262 UV-endonucle UV-species: N C; Species: 10-C C; Accession: R; Yajima, H. R; Yajima, H. R; Yajima, H. R; Yajima, H. R; Azitle: A e A; Reference A; Reference A; Reference A; Reference A; Reference A; Reseaule t A; Molecule t A; Molecule t A; Cossionics: A; Residues: A; Crossiones: A; Crossiones: A; Crossiones: A; Crossiones: A; Genetics:	Query Mat Best Loca Matches	QY Db	Qy	Qy 1 Db 1	Qy 1 Db 1	Qy 2 Db 2	Qy 3	QY 31
	4			,				

421 ENDDVGWTVHDLLPVCEELNIPMVLDYHHHNICFDPAHLREGTLDISDPKLQERIANTWK 480

QQ ŏ



QY 354 EVVESAIRDLEYHDELLSLLKLPEQQNRDAVMIIHMGGOFGDKAATLERFKRNYARLSQS 413	Db 543 VFELCRRYEL	RESULT 3 E83974 hypothetical protein BH2597 [imported] - Bacillus halodurans (strain C-125) C:Species: Bacillus halodurans C:Decies: Dace-2000 #sequence_revision 01-Dec-2000 #text_change 31-Dec-2000 C:Accession: E83974 R:Takami, H: Nakasone, K:; Takaki, Y:; Maeno, G:; Sasaki, R.; Masui, N.; Fuji, F:; H Nucleic Acids Res. 28, 4317-4331, 2000	A; Reference number: A83650; MUID:20263314 A; Reference number: A83650; MUID:20263314 A; Reference number: A83650; MUID:20263314 A; Reference number: R83974 A; Status; preliminary A; Molecule type: DAA A; Residues: 1-322 <sto> A; Cross-references: GB:AP001516; GB:BA000004; NID:g10175192; PIDN:BAB06316.1; GSPDB:G</sto>		OY 207 RIGYACLNTYLRNARPPIFSSRTCRMASIVDHRHPLQFEDEPEHHLKNKPDKSKEPQDEL 266			
Db 421 ENDDVGWTVHDLLPVCEELNIPHVLDYHHNICFDPAHLREGTLDISDPKLGERIANTWK 480 Qy 481 RKGIKOKMHYSEPCDGAVTPRDRRKHRPRVWTLPPCPPDMDLMIEAKDKEQAVFELMRTF Db 481 RKGIKOKMHYSEPCDGAVTPRRRKHRPRVWTLPPCPPDMDLMIEAKDKEQAVFELMRTF Qy 541 KLPGFEKINDWVPYDRDDENRPAPPVKAPKKKGGKRKRTTDERAAEPEEVDTAADDVKD Db 541 KLPGFEKINDWVPYDRDDENRPAPPVKAPKKKGGKRKRTTDERAAEPEEVDTAADDVKD OY 601 APEGPKEVPEERAAMGGPYNRYWWPLGCEEWLKPKKREVKKGVPEEVEDGGEFG 656 Db 601 APEGPKEVPEERAAMGGPYNRYYWPLGCEEWLKPKKREVKKGKVPEEVEDGGEFG 656 Db 601 APEGPKEVPEERAAMGGPYNRYYWPLGCEEWLKPKKREVKKGKVPEEVEDGGEFG 656	RESULT 2 S71134 UV-endonuclease - fission yeast (Schizosaccharomyces pombe) Species: Schizosaccharomyces pombe hte: 21-Apr-1997 #sequence_revision 09-May-1997 #text_change 20-Jun-2000 Accession: S71134: T39815	R:Takao, M.; Yonemasu, R.; Yamamoto, K.; Yasui, A. Nucleic Acids Res. 24, 1267-1271, 1996 A:Title: Characterization of a UV endonuclease gene from the fisson yeast Schizosaccharc A:Reference number: S7134; MUID:96188860 A:Accession: S7134 A:Accession: S7134 A:Accession: S7134 A:Molecule type: mRNA A:Residues: 1-599 <	1	A:Experimental source: strain 972h-; cosmid c19c7 C;Genetics: A:Gene: UVDE; SPBC19c7.09c A:Map position: 2	Query Me Best Loc tches	110 63 158	115 185 174 233	OY 234 SIVDHRHPLQFEDEPEHHLKNKPDKSKEPQDELGHKFVQELGLANARDIVKMLCWNEKYG 293 :



```
(without alignments)
630.625 Million cell updates/sec
                                                                                                      ; Search time 134.53 Seconds
                                                                                                                                                                                                          1 MGTTGLESLSLGDRGAAPTV......HKLNKYHDWLWENHEKLSLS
                                                                                                                                                                                                                                                                                                                                          473505
GenCore version 4.5
Copyright (c) 1993 - 2000 Compugen Ltd.
                                                                                                                                                                                                                                                                                                                                 Total number of hits satisfying chosen parameters:
                                                                                                                                                                                                                                                                                                     473505 seqs, 146272329 residues
                                                                                                                                                                                                                                                                                                                                                                                                                            Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries
                                                                                                    January 15, 2002, 14:06:28
                                                                  OM protein - protein search, using sw model
                                                                                                                                                                                                                                                                  Gapop 10.0 , Gapext 0.5
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      sp_virus:*
sp_vertebrate:*
sp_unclassified:*
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           sp_invertebrate:*
sp_mammal:*
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            sp_organelle:*
sp_phage:*
                                                                                                                                                                                                                                                                                                                                                                        Minimum DB seq length: 0
Maximum DB seq length: 200000000
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     sp_archea:*
sp_bacteria:*
                                                                                                                                                                            US-09-724-296-38
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   sp_rodent:*
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     sp_plant: *
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          sp_fungi:*
sp_human:*
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    SPTREMBL_17:*
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             sp_mhc:*
                                                                                                                                                                                                                                                   BLOSUM62
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     10:
                                                                                                                                                                                               Perfect score:
                                                                                                                                                                                                                                                 Scoring table:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    Database :
                                                                                                                                                                                                                                                                                                     Searched:
                                                                                                                                                                                                               Sedneuce:
                                                                                                      Run on:
                                                                                                                                                                            Title:
```

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

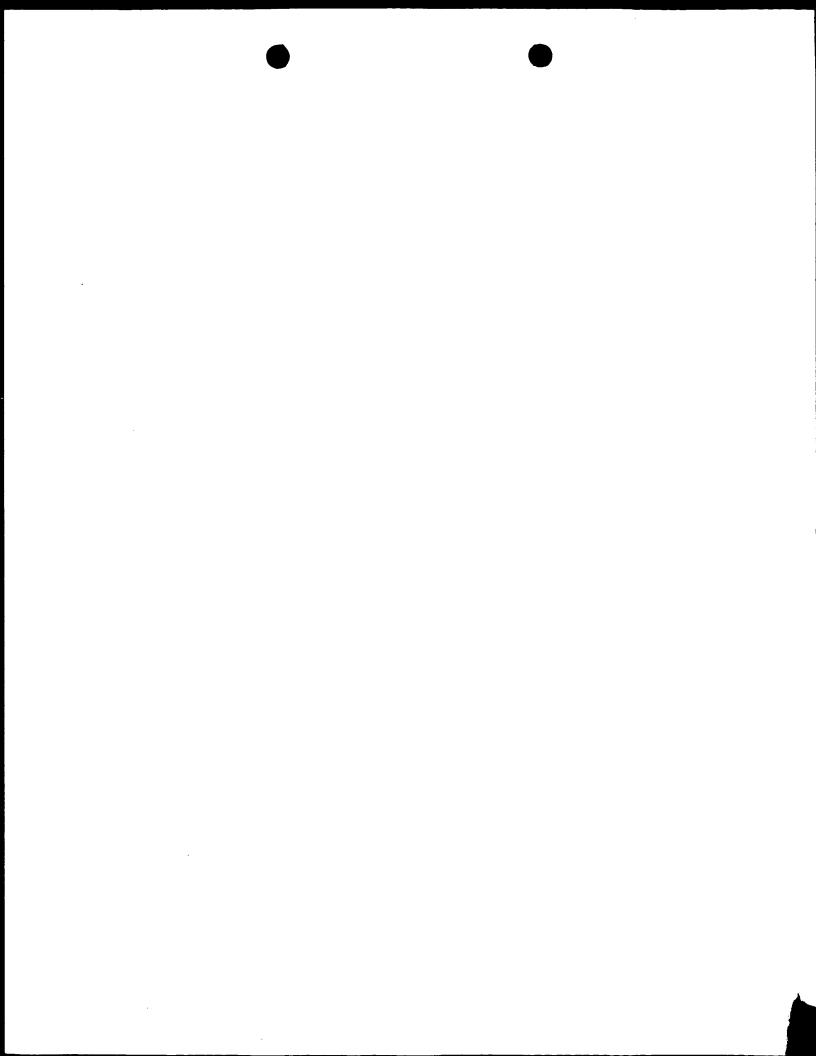
SUMMARIES

					SUMMARIES	
Result		Ouery				
NO.	Score	Match	Match Length DB	DB	ID	Description
-	3055	100.0	580	4	095243	095243 homo sapien
7	1821.5	59.6	554	11	Q9Z2D7	09z2d7 mus musculu
٣	880.5	28.8	416	13	Q919F1	09i9fl gallus gall
4	344	11.3	419	10	Q9SFC1	
ഗ	220.5	7.2	467	13	Q9YGC6	
9	202.5	9.9	344	13	042403	042403 gallus gall
7	146	4.8	702	10	Q9SF22	09sf22 arabidopsis
80	143.5	4.7	782	S	025875	025875 plasmodium
σ	141.5	4.6	782	S	025730	
10	141.5	4.6	782	Ŋ	026007	O26007 plasmodium
11	140.5	4.6	782	വ	Q9U430	Ogu430 plasmodium
12	140.5	4.6	782	Ŋ	090414	Ogu414 plasmodium
13	138.5	4.5	782	ა	090429	Q9u429 plasmodium
14	137.5	4.5	782	Ŋ	026104	Q26104 plasmodium
15	137.5	4.5	2209	S	900060	Q9u0q6 plasmodium
16	136.5	4.5	782	Ŋ	090431	O9u431 plasmodium
17	136.5	4.5	1359	Ŋ	002061	002061 caenorhabdi
18	135.5	4.4	282	13	Q9PUM9	Q9pum9 xenopus lae
19	135.5	4.4	755	4	043719	043719 homo sapien

20 135.5 4.4 1058 5 09V433 21 135 4.4 822 3 09U5E0 23 133.5 4.4 1359 5 09U7E0 24 132.5 4.3 1819 2 09U7E0 24 131.5 4.3 1819 2 09ZLV0 25 131.5 4.3 285 11 09PDL6 27 131.5 4.3 29L 4 09598 20 130.5 4.3 1819 2 09ZLV0 29 130.5 4.3 29L 4 09598 30 130.5 4.3 1927 2 02526 31 129.5 4.2 1163 4 09UBS 33 129.5 4.2 1163 4 09UBS 34 128.5 4.2 168 5 09UE0 83 127 4.2 168 5 09UE0 84 125.5 4.1 183 2 09BUMB 41 125.5 4.1 302 4 09USE0 41 125.5 4 11 302 4 09USE0 41 125.5 4 11 302 4 125.5 4 125.5	Q9v433 drosophila	Q9ush9 schizosacch	Q9u7e0 caenorhabdi	Q9ydp0 aeropyrum p	09zlv0 helicobacte	09p916 pyrobaculum	O9z2d8 mus musculu	095983 homo sapien		Q9ubb5 homo sapien	Q9vzt4 drosophila	025262 helicobacte	081mge mns mnscnin	Ogunb7 homo sapien	09c9d8 arabidopsis	091fe4 arabidopsis	Q9u5e0 drosophila	091qv6 arabidopsis	Q9ayf2 oryza sativ	Q9jmx8 helicobacte	Q9w3dl drosophila	Q9pum8 xenopus lae	09c6s8 arabidopsis	Q9bkn8 plasmodium	Q9uis8 homo sapien	Q9wvc9 mus musculu
135.5 4.4 1058 133.5 4.4 1359 132.5 4.4 1359 132.5 4.3 1819 131.5 4.3 230 131.5 4.3 231 131.5 4.3 291 131.5 4.3 291 130.5 4.3 291 130. 4.3 1927 130. 4.2 3122 129.5 4.2 3122 128.5 4.2 946 127.5 4.2 1058 127.5 4.2 970 127.6 4.1 1033 126.5 4.1 1033	Q9V433	6HSD6O	Q9U7E0	Q9YDP0	09ZLV0	Q9P9L6	Q922D8	095983	099730	Q9UBB5	Q9VZT4	025262	90MD60	Q9UHB7	Q9C9D8	Q9LFE4	Q905E0	9A0760	Q9AYF2	о9лмх8	Q9W3D1	09PUM8	090658	Q9BKN8	851060	Q9WVC9
135.5 133.5	Ŋ	٣	Ŋ	-	~	Н	Π	7	4	4	'n	7	11	4	10	10	Ŋ	10	10	7	2	13	10	Ŋ	4	11
135.5 133.5 133.5 133.5 131.5 131.5 131.5 131.5 130.5 129.5 129.5 129.5 127.5 127.5 127.5 126.5 126.5 126.5 126.5 127.5	1058	822	1359	286	1819	230	285	291	754	411	537	1927	3122	1163	946	853	1058	648	970	1723	1183	303	561	881	302	873
	4.4	4.4	4.4	4.3	4.3	4.3	4.3	4.3	4.3	4.3	4.3	4.3	4.2	4.2	4.2	4.2	4.2	4.2	4.2	4.2	4.1	4.1	4.1	4.1	4.1	4.1
04040404040404040404040404040404040404	135.5	135	133.5	132.5	132	131.5	131.5	131.5	131	130.5	130	130	129.5	129	128.5	128	127.5	127	127	127	126.5	126	126	126	125.5	125
	20	21	22	23	24	25	56	27	28	29	30	31	32	33	34	35	36	37	38	39	40	41	42	43	44	45

ALIGNMENTS

```
SEQUENCE FROM N.A. MEDLINE=98449942; PubMed=9774669; Hendrich B., Bird A.; "Identification and characterization of a family of mammalian methyl-
                                                                                                                                                                                                                                                                                                                                                 TISSUE-FETAL BRAIN;
MEDLINE=99199294; PubMed=10097147;
Bellacosa A., Gicchillitti L., Schepis F., Riccio A., Yeung A.T.,
Matsumoto Y., Golemis E.A., Genvardi M., Neri G.;
Matsumoto Y., Golemis E.A., Genvardi M., Neri G.;
DNA mismatch repair protein MLHI.";
Proc. Natl. Acad. Sci. U.S.A. 96:3969-3974(1999).
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            MEDLINE-9937225; PubMed=10441743; Hendrich B., Abbott C., McQueen H., Chambers D., Cross S., Bird A.; "Genomic structure and chromosomal mapping of the murine and human
                                                                                                                                                        Eukaryota, Metazoa; Chordata: Craniata; Vertebrata; Buteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
                                                           01-WAY-1999 (TrEMBLrel. 10, Created)
01-WAY-1999 (TrEMBLrel. 10, Last sequence update)
01-JUN-2001 (TrEMBLrel. 17, Last annotation update)
                                  580 AA.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            mbd1, mbd2, mbd3, and mbd4 genes.";
Magm. Genome 10:906-912(1999).
EMBL; AF072250; AAC68879.1; -
EMBL; AF114784; AAD22195.1; -
EMBL; AF12099; AAD50374.1; JOINED.
EMBL; AF12099; AAD50374.1; JOINED.
                                                                                                                                                                                                                                                                                    CpG binding proteins.";
Mol. Cell. Biol. 18:6538-6547(1998).
                                PRT;
                                                                                                          METHYL-CPG BINDING PROTEIN MBD4.
                                PRELIMINARY;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            InterPro; IPR001739; MBD
                                                                                                                                           Homo sapiens (Human).
                                                                                                                                                                                                                                                                                                                                                                                                                                                                               SEQUENCE FROM N.A.
                                                                                                                                                                                                                                                                                                                          SEQUENCE FROM N.A.
                                                                                                                                                                                         NCBI_TaxID=9606;
                                                                                                                            MBD4 OR MED1.
                                              095243;
                               095243
RESULT
                095243
```

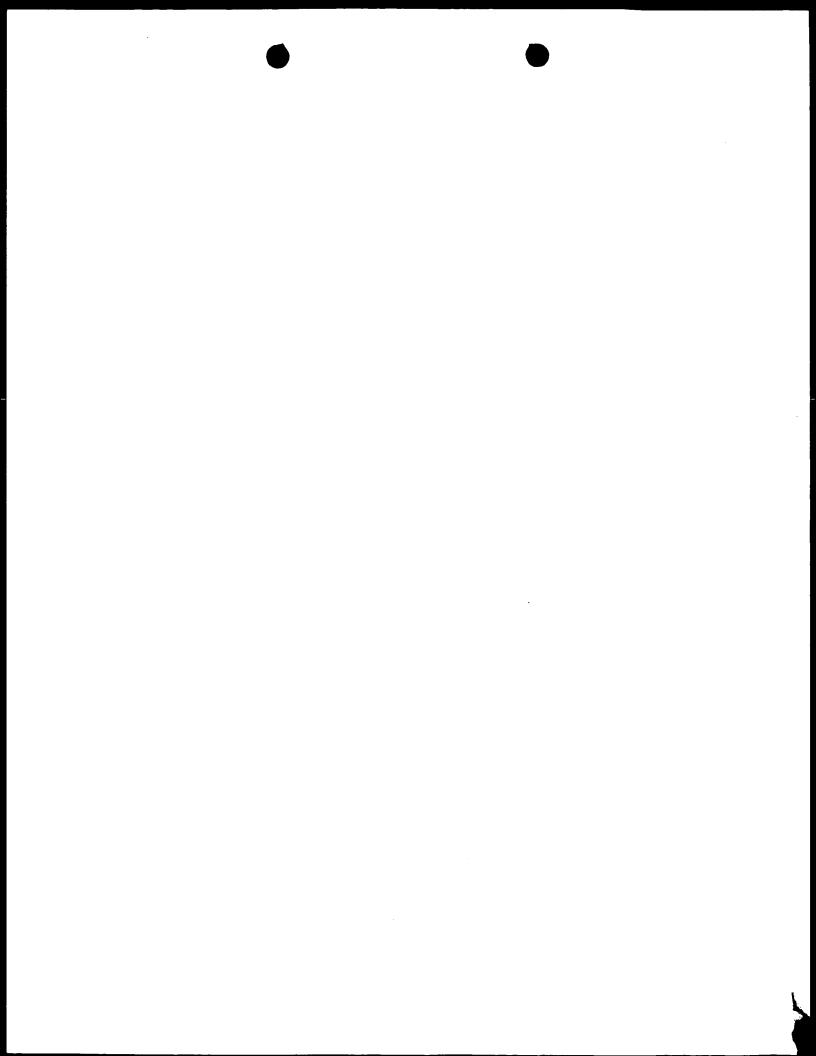


a

Page

```
Query Match
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    Q919F1
Q919F1;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  RESULT 3
                                                                                                                                                                                                                                                                                                                                                                                           110
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                515
                                                                                                                                                                                                                                                                                                                                                                      123
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            0919F1
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       FIRE
                                                                                                                                                                                                                                                                                                      qq
                                                                                                                                                                                                                                                                                                                           δ
                                                                                                                                                                                                                                                                                                                                                                                           q
                                                                                                                                                                                                                                                                                                                                                                                                                                   QQ
                                                                                                                                                                                                                                                                                                                                                                                                                                                                              qq
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    à
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        qq
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  q
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        δy
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            qq
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   δ
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     q
  ô
                                                                                                                                                                                                                                                                                                                                                                                                                ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                          Q_{Y}
                                                                                                                                                                                                                                                                                  ij.
                                                                                                           ö
                                                                                                                                                                                                                  180
                                                                                                                                                                                  61 CNPLLQEPIASAQFGATAGTECRKSVPCGWERVVKQRLFGKTAGRFDVYF1SPQGLKFRS 120
                                                                                                                                                                                                                          121 KSSLANYLHKNGETSLKPEDFDFTVLSKRGIKSRYKDCSMAALTSHLONGSNNSNWNLRT 180
                                                                                                                                                                                                                                                                                                     241 IPIKKTKKGCRKSCSGFVQSDSKRESVCNKADAESEPVAQKSQLDRTVCISDAGACGETL 300
                                                                                                                                                                                                                                                                                                               360
                                                                                                                                                                                                                                                                                                                                                                     360
                                                                                                                                                                                                                                                                                                                                                                                                                420
                                                                                                                                                                                                                                                                                                                                                                                                                                    480
                                                                                                                                                                                                                                                                                                                                                                                                                                                                               540
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         SEGUENCE FROM N.A.
MEDLINE-98449942; PubMed-9774669;
Hendrich B., Bird A.;
"Identification and characterization of a family of mammalian methyl-
                                                                                                                                                                        61 CNPLLQEPIASAQFGATAGTECRKSVPCGWERVVKQRLFGKTAGRFDVYFISPQGLKFRS 120
                                                                                                          Gaps
                                                                                                                                        1 MGTTGLESLSLGDRGAAPTVTSSERLVPDPPNDLRKEDVAMELERVGEDEEQMMIKRSSE 60
                                                                                                                                                                                                                                                                        121 KSSLANYLHKNGETSLKPEDFDFTVLSKRGIKSRYKDCSMAALTSHLQNQSNNSNWNLRT
                                                                                                                                                                                                                                                           RSKCKKDVFMPPSSSSELQESRGLSNFTSTHLLLKEDEGVDDVNFRKVRKPKGKVT1LKG
                                                                                                                                                                                                                                                                                                                                                SVTSEENSLVKKKERSLSSGSNFCSEQKTSGIINKFCSAKDSEHNEKYEDTFLESEEIGT
                                                                                                                                                                                                                                                                                                                                                          KYNKEALSPPRRKAFKKWTPPRSPFNLVQETLFHDPWKLLIATIFLNRTSGKMAIPVLWK
                                                                                                                                                                                                                                                                                                                                                                                                                                                                            481 FLEKYPSAEVARTADWRDVSELLKPLGLYDLRAKTIVKFSDEYLTKQWKYPIELHGIGKY
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus
NCBL_TaxID=10090, 10092;
                                                                                                                                                                                                                                                                                                                                                                                         KVEVVERKEHLHTDILKRGSEMDNNCSPTRKDFTGEKIFQEDTIPRTQIERRKTSLYFSS
                                                                                                          .
0
                                                                                   100.0%; Score 3055; DB 4; Length 580; 100.0%; Pred. No. 1.9e-219; ive 0; Mismatches 0; Indels 0
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     Mus musculus domesticus (western European house mouse)
                                                     BF16FB21A34B8E5F CRC64;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   01-MAY-1999 (TrEMBLrel. 10, Created)
01-MAY-1999 (TrEMBLrel. 10, Last sequence update)
01-JUN-2001 (TrEMBLrel. 17, Last annotation update)
METHYL-CPG BINDING PROTEIN MBD4.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       541 GNDSYRIFCVNEWKQVHPEDHKLNKYHDWLWENHEKLSLS 580
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   541 GNDSYRIFCVNEWKQVHPEDHKLNKYHDWLWENHEKLSLS 580
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 554 AA
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 PRT;
                                                     580 AA; 66050 MW;
InterPro; IPR003265; Endo_3c.
         Pfam; PF00730; HhH-GPD; 1.
Pfam; PF01429; MBD; 1.
SMART; SM00391; MBD; 1.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         Mus musculus (Mouse), and
                                                                                                       580; Conservative
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 PRELIMINARY;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      CpG binding proteins.";
                                                                                              Best Local Similarity
                                          Endonuclease
                                                     SEQUENCE
                                                                                     Query Match
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                Q922D7
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           0922D7
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            7
                                                                                                                                                                                                                                                           181
                                                                                                                                                                                                                                                                                                                                                                    301
                                                                                                                                                                                                                                                                                                                                                                                                                                   421
                                                                                                          Matches
                                                                                                                                                                                                                                                                                                                                               301
                                                                                                                                                                                                                                                                                                                                                                                         361
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            RESULT
Q922D7
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                g
                                                                                                                                                                                                                 ò
                                                                                                                                                                                                                                      a
                                                                                                                                                                                                                                                                            g
                                                                                                                                                                                                                                                                                                                     g
                                                                                                                                                                                                                                                                                                                                         õ
                                                                                                                                                                                                                                                                                                                                                                엄
                                                                                                                                                                                                                                                                                                                                                                                                         g
                                                                                                                                                                                                                                                                                                                                                                                                                                 ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                      g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       ö
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           셤
ò
                                                                                                                                                                        ò
                                                                                                                                                                                                                                                           ò
                                                                                                                                                                                                                                                                                                    ò
                                                                                                                                                                                                                                                                                                                                                                                        οy
                                                                                                                                                                                                                                                                                                                                                                                                                                                                              à
```

```
363 EVVERK -- EHLHTDILKRGSEMDNNCSPTRKDFTGEKIFQEDTIPRTQIERRKTSLYFSS 420
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    421 KYNKEALSPPRRKAFKKWTPPRSPFNLVQETLFHDPWKLLIATIFLNRTSGKMAIPVLWK 480
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            PLLQEPIASAQFGATAGTECRKSVPCGWERVVKQRLFGKTAGRFDVYFISPQGLKFRSKS 122
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           SLANYLHKNGETSLKPEDFDFTVLSKRGIKSRYKDCSMAALTSHLQNQSNNSNWNLRTRS 182
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            SLANYLLKNGETFLKPEDFNFTVLPKGSINPGYKHOSLAALTSLQPNETDVSKONLKTRS 169
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           303 TSEENSLVKKKERSLSSGSNFCSEQKTSGIINKFCSAKDSEHNEKYEDTFLESEEIGTKV 362
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               Gaps
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                6 LESLSLGD----RGAAPTVTSSERLVPDPPNDLRKEDVAMELERVGEDEEQMMIKRSSECN 62
                                                                                                                                  Bird A.;
                                                                                                                                                       mapping of the murine and human
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                243 IKKTKKGCRKSCSGFVQSDSKRESVCNKADAESEPVAQKSQLDRTVCISDAGACGETLSV
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  183 KCKKDVFMPPSSSSELQESRGLSNFTSTHLLLKEDEGVDDVNFRKVRKPKGKVTILKGIP
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      481 FLEKYPSAEVARTADWRDVSELLKPLGLYDLRAKTIVKFSDEYLTKQWKYPIELHGIGKY
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      59.6%; Score 1821.5; DB 11; Length 554; 66.2%; Pred. No. 1.4e-127; Live 49; Mismatches 116; Indels 31;
                                                                                                                               Chambers D., Cross S.,
                                                                                                                                                                                                                                                                                                                                                                                                                                                            PROSITE; PS00092; N6_MTASE; UNKNOWN_1.
SEQUENCE 554 AA; 62577 MW; 792D37CB180291F5 CRC64;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            541 GNDSYRIFCVNEWKQVHPEDHKLNKYHDWLWENHEKLSLS 580
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                01-0CT-2000 (TrEMBLrel. 15, Created)
01-0CT-2000 (TrEMBLrel. 15, Last sequence update)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        Cell. Biol. 18:6538-6547(1998).
                                                                                                                            Hendrich B., Abbott C., McQueen H., "Genomic structure and chromosomal mmbd1, mbd2, mbd3, and mbd4 genes.";
                                                                                                    MEDLINE=99373255; PubMed=10441743;
                                                                                                                                                                                                     Mamm. Genome 10:906-912(1999).
EMBL; AF072249, AAC68878.1; --
EMBL; AF120996; AAD56595.1; --
MGD; MGI:1333850; MD44.
                                                                                                                                                                                                                                                                                                             InterPro; IPR003265; Endo_3c.
InterPro; IPR001739; MBD.
InterPro; IPR002052; N6_Mtase.
                                                                                                                                                                                                                                                                                                                                                                                        Pfam; PF00730; HhH-GPD; 1.
Pfam; PF01429; MBD; 1.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              Best Local Similarity 66.2
Matches 384; Conservative
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     PRELIMINARY;
                                                                                                                                                                                                                                                                                                                                                                                                                                               SM00391; MBD;
                                                   SEQUENCE FROM N.A.
```



```
GenCore version 4.5
Copyright (c) 1993 - 2000 Compugen Ltd.
                                                                                                  OM protein - protein search, using sw model
```

(without alignments) 841.640 Million cell updates/sec Search time 74.94 Seconds January 15, 2002, 13:45:51 US-09-724-296-2 Perfect score: Run on: Title:

1 MTKLPILGYWKIKGLVQPTR........KRLTARKRRSRKEEVEEDEK 828 **BLOSUM62** Scoring table: Sequence:

Gapop 10.0 , Gapext 0.5

219241 seqs, 76174552 residues Searched:

219241 Total number of hits satisfying chosen parameters: Minimum DB seq length: 0 Maximum DB seq length: 200000000

Post-processing: Minimum Match 0% Maximum Match 100% Listing first 45 summaries

pir1:* pir2:* pir3:* PIR_68:* Database

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

glutathione transf glutathione S-tran glutathione transf glutathione transf glutathione transf glutathione transf glutathione S-tran UV-endonucléase ho hypothetical prote glutathione transf UV-endonuclease transf transf transf transf transf mu-class glutathio glutathione transf transf transf transf transf transf transf transf transf glutathione transf major allerqen Dp1 UV-endonuclease Description glutathione SUMMARIES A47486 B34159 S01719 A46048 A29036 A46143 B28946 A48388 S33860 S32425 A45556 A29794 865674 JX0095 A39375 \$13202 B29231 XURTG4 A35295 S18464 355418 S17462 E83974 Match Length DB 22.6 21.2 15.0 11.1 11.1 11.0 10.9 10.9 10.9 10.8 Query 10.7 10.6 10.6 10.5 Score 481.5 481.5 478.5 477.5 474.5 473.5 473.5 470 468.5 465.5 463.5 461.5 460.5 455.5 467.5 452.5 360 306.5 385. Result

(
gracuitone cransi	,	1	•			
	203615	0	208	ر د	219.5	S
dlutathione transf	441031] A	223	5.1	5.777	4
probable UV damage	73330) 7	320) i	0.000) •
	0000	,	308	رب س	229 5	~
*	A49180	2 A	210	5.3	230	2
alutathione transf	441177	2 A	210		232	-
glutathione transf	455140	T A	507		233	٠,
qlutathione transf	706529	7	710		0.00	n 0
glutathione transf	5.1.9.3	9 (0 4 6		2 0	
	04017		010		236	8
	148112	2 I	210	5.4	236	37
alutathione transf	355140	I H	210		757	0 1
glutathione transf	URIGE	ς -	777		- 1	1 (
	000000000000000000000000000000000000000	,	010		237	S.
-	371957	C1	210		238	34
glutathione transf	A37378	₹.	210		0 4 7	η.
glutathione transf	13/80	7	202		4 0	10
_	00000		202	V.	243	3.2
	371958	2	210	5.6	244	31
alutathione transf	S41933	2	208	5.6	245.5	30

UV-endonuclease - fission yeast (Schizosaccharomyces pombe)
C;Species: Schizosaccharomyces pombe
C;Species: Schizosaccharomyces pombe
C;Becies: 21-Apr. 1997 #sequence_revision 09-May-1997 #text_change 20-Jun-2000
C;Accession: S71134; T39815
R;Takao, M.; Yonemasu, R.; Yamamoto, K.; Yasui, A.
Nucleic Acids Res. 24, 1267-1271, 1996
ArTitle: Characterization-of-active Admension Programment Artitle: Characterization-of-active Acids Res. 24, 1267-1271, 1996
Artitle: Characterization-of-active Acids Res. Schizosacch

RESULT

lekao

A; Status: preliminary; nucleic acid sequence not shown A; Reference number: S71134; MUID:96188860 A; Accession: S71134

A; Molecule type: mRNA A; Residues: 1-599 <TAK>

A)Cross-references: EMBL:D78571; NID:g1399000; PIDN:BAA11415.1; PID:g1399001 A)Experimental source: strain SP972 R:Lyne, M.; Rajandream, M.A.: Barrell, B.G.: Lucas, M.; Gaillardin, C. submitted to the EMBL Data Library, June 1998 A)Reference number: Z21881 A)Accession: T39815

A;Status: preliminary; translated from GB/EMBL/DDBJA:Molecule type: DNA

A;Residues: 1-599 <LYN> A;Cross.references: EMBL:AL023859; PIDN:CAA19577.1; GSPDB:GN00067; SPDB:SPRC19C7.09c A;Experimental source: strain 972h-; cosmid c19C7 A; Gene: UVDE; SPBC19C7.09c C; Genetics:

A; Map position: 2

230 MLRLLKRNIQISKRIVFTILKQKAFKGNHPCVPSVCTITYSRFHCLPDTLKSLLPMSSKT 289 290 TLSMLPQVNIGANSFSAETPVDLKKENETELANISGPHKKSTSTSTRKRARSSKKKATDS 349 350 VSDKIDESVASYDSSTHLRRSSRSKKPVNYNSSSESEEGISKATKKVKQKEEEETVEE 409 410 VDEKSLKNESSSDEFEPVVPEQLETPISKRRRSRSSAKNLEKESTMNLDDHAPREMFDCL 469 0; Gaps 71.6%; Score 3112; DB 2; Length 599; 100.0%; Pred. No. 1.8e-162; Live 0; Mismatches 0; Indels Indels Conservative Query Match Best Local Similarity Matches 599; QQ δy qq δ qq òγ

ö

qq

DKPIPWRGRLGYACLNTILRSMKERVFCSRTCRITTIQRDGLESVKQLGTQNVLDLIKLV 529 470 ΟŊ

18;

Gaps

```
A;Title: A eukaryotic gene encoding an endonuclease that specifically repairs DNA dam A;Reference number: S55262; MUID:95292980
A;Accession: S55262
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        Rivight, M.D.; Harrison, R.A.; Melder, A.M.; Newport, G.R.; Mitchell, G.F. Mol. Biochem. Parasitol. 49, 177-179, 1991
Mol. Biochem. Parasitol. 49, 177-179, 1991
Affile: Another. 26-kilodalton glutathione S-transferase of Schistosoma mansoni. A;Reference number: A45556; MUID:92131046
A;Accession: A45556
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               glutathione S-transferase - fluke (Schistosoma mansoni)
C;Species: Schistosoma mansoni
C;Date: 22-Apr-1993 #sequence_revision 18-Nov-1994 #text_change 03-May-1996
C;Accession: A45556
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           596 VVVDSAIRDLAYHDEILSRMKLNEQLNKDAVLIIHLGGTFEGKKETLDRFFKNYORLSDS 655
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     656 VKARLVLENDDVSWSVQDELPLCQEENIPLVLDWHHHNIV--PGTLREGSLDLM--PLIP 711
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               712 TIRETWTRKGITQKQHYSESADPTAISGMKRRAHSDRVFDFPPCDPTMDLMIEAKEKEQA 771
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     772 VFELCRRYEL-----LEIMGPE--- 794
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                533 VFELMRTFKL.PGFEKINDMVPYDRDDENRPAPPVKAPKKKGGKRKRTTDEEAAEPEEVD 592
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                234 SIVDHRHPLOFEDEPEHHLKNKPDKSKEPQDELGHKFVQELGLANARDIVKMLCWNEKYG 293
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       537 IHFMRVSSDLFPFASHAKYGYTL-EFAQSHLEEVGKLANKYNHRLTMHPGQYTQIASPRE 595
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  294 IRFLRLSSEMFPFASHPVHGYKLAPFASEVLAEAGRVAAELGHRLTTHPGQFTQLGSPRK 353
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                339 ARSSKKKATDSVSDKIDESVASYDSST-----HLRRSSR-----SKKPVNYNSSSESE- 386
                                                                                                                                                                                                                                                                                                                                                                                                                                       387 -----SEEQISKATKKVKQKEEEEYVEEVDEK 413
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          795 --YDQTRDG-----EVEED 826
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        593 TAADDVKDAPEGPKEVPEEERAMGGPYNRVYWPLGCEEWLKPKKREVKKGKVPEEVEDE 651
                                                                                                                                                                                                                                                                                                                                                                                            3 SRKSKAAALDTPQSESSTFSSTLDSSAPSPARNLRRSGRN1LQPSSEKDRDHEKRSGEEL 62
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        462 ----PREMFDCLDK------PIPWRGRLGYACLNTILRSMKERVFCSRTCRIT
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   ------QRDGLESVKQLGTQNVLDLIKLVEWNHNFG
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       63 AGRMMGKDANGHCLREGKEQEEGVKMAIEGLARMERRLQRATKRQKKQLEED-----
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     414 SLKNESSSDEFEPVVPEQLETPISKRRRS-----RSSAKNLEKESTMNLDDHA-----
                                                                                                                                                                                                                                                      23.3%; Score 1011.5; DB 2; Length 656; 36.6%; Pred. No. 7.6e-48; Live 88; Mismatches 149; Indels 181;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           A;Status: preliminary; not compared with conceptual translation A;Molecule type: nucleic acid
                                                                                                                  A;Residues: 1-656 <VAJ>
A;Cross-references: EMBL:D11392; NID:q1526560; PID:g927215
C;Genetics:
                                                                            A;Status: nucleic acid sequence not shown
                                                                                                                                                                                                                                                                                                             Matches 241; Conservative
                                                                                                                                                                                                                                                                  Query Match
Best Local Similarity
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             205 TI-----
                                                                                                      A; Molecule type: mRNA
                                                                                                                                                                                                    A; Gene: uvel
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       4
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     RESULT
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             A45556
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         Óλ
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    q
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             Db
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           ŏ
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   qq
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 δy
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         q
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            QQ
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       Ob
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  QΥ
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                q
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             ōλ
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      qq
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      δŏ
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            δy
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 qq
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            ŏ
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       qq
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    C; Species: Schistosoma japonicum
C; Species: Schistosoma japonicum
C; Date: 25-Oct-1987 #sequence_revision 30-Sep-1989 #text_change 03-Feb-1994
C; Accession: A94139; A26444; A28315
R; Smith, D. B.; Davern, K.M.; Board, P.G.; Tiu, W.U.; Garcia, E.G.; Mitchell, G.F.
R; Smith, D. B.; Davern, K.M.; Board, P.G.; Tiu, W.U.; Garcia, E.G.; Mitchell, G.F.
A; Title: Mr 26,000 antigen of Schistosoma japonicum recognized by resistant WEHI 129/J
A; Accession: A94139
A; Molecule type: mRNA
A; Molecule type: mRNA
A; Molecule type: M. M.; Board, P.G.; Tiu, W.U.; Garcia, E.G.; Mitchell, G.F.
Proc. Natl. Acad. Sci. U.S.A. 84, 6541, 1987
A; Reference number: A94181
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      C;Species: Neurospora crassa
C;Date: 10-Oct-1995 #sequence_revision 03-Nov-1995 #text_change 20-Jun-2000
C;Accession: S55262
R;Yajima, H.; Takao, M.; Yasuhira, S.; Zhao, J.H.; Ishii, C.; Inoue, H.; Yasui, A.
EMBO J. 14, 2393-2399, 1995
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            glutathione transferase (EC 2.5.1.18) - fluke (Schistosoma japonicum) (fragment)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       0
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 VDFLSKLPEMLKMFEDRLCHKTYLNGDHVTHPDFMLYDALDVVLYMDPMCLDAFPKLVCF 181
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            62 DGDVKL,TQSMAIIRYIADKHNMLGGCPKERAEISMLEGAVLDIRYGVSRIAYSKDFETLK 121
                                                                                                                                                                                                                                                                       710 IPTIRETWTRKGITQKQHYSESADPTAISGMKRRAHSDRVFDFPPCDPTWDLMIEAKEKE 769
                                                                                                                                                                                                                                                                                                                                                               QRLSDSVKARLVLENDDVSWSVQDLLPLCQELNIPLVLDWHHHNIVPGTLREGSLDLMPL 709
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                2 TKLPILGYWKIKGLVOPTRLLLEYLEEKYEEHLYERDEGDKWRNKKFELGLEFPNLPYYI 61
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          1 TKLPILGYWKIKGLVQPTRLLLEYLEEKYEEHLYERDEGDKWRNKKFELGLEFPNLPYYI 60
                241 DKPIPWRGRLGYACLNTILRSMKERVFCSRTCRITTIQRDGLESVKQLGTQNVLDLIKLV 300
                                                                EWNHNFGIHFMRVSSDLFPFASHAKYGYTLEFAQSHLEEVGKLANKYNHRLTMHPGQYTQ 589
                                                                                                                                                         590 IASPREVVVDSAIRDLAYHDEILSRMKLNEQLNKDAVLIIHLGGTFEGKKETLDRFRKNY 649
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            Gaps
                                                                                                                                                                                                                                                                                                                                                                                                                                770 QAVFELCRRYELQNPPCPLEIMGPEYDQTRDGYYPPGAEKRLTARKRRSRKEEVEEDEK 828
                                                                                                                                                                                                                                                                                                                                                                                                                                                        0;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               Length 219;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            Indels
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            27.2%; Score 1184; DB 2;
100.0%; Pred. No. 6.7e-58;
tive 0; Mismatches 0;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  182 KKRIEAIPQIDKYLKSSKYIAWPLQGWQATFGGGDHPPK 220
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              A;Contents: annotation; revision to residues 210-219 C;Superfamily: glutathione transferase C;Keywords: transferase
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          UV-endonuclease - Neurospora crassa
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         Best Local Similarity 100.
Matches 219; Conservative
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      Query Match
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           122
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 Matches
                                                                                                                     301
                                                                                                                                                                                                                                                         650
                                                                       530
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       S55262
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    QQ
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              ŏ
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  δ
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             q
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           ó
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   ŏ
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            g
                                                                                                                                                                                                           q
                                                                                                                                                                                                                                                                                                    QO
                                                                                                                                                                                                                                                                                                                                                 δy
                                                                                                                                                                                                                                                                                                                                                                                               qq
                                                                                                                                                                                                                                                                                                                                                                                                                                            ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                Op
                                                                                                                     g
                                                                                                                                                                δ
                                                                                                                                                                                                                                                         δ
```

536

461

```
A; Experimental source: Puerto Rico
                                                                                                                                                                                                                                                                                                                                                                                                                                                                   15.0%;
58.2%;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      Matches 121; Conservative
                                                                                                                                                                                                                                                                                           A;Status: preliminary
A;Molecule type: nucleic acid
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       Best Local Similarity
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         A; Contents: clone pGT 875
                                                                                                                                                                                                                                                                                                                                       -209 <MUR>
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    DNA
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 A; Molecule type: mRNA
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            833860
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               A;Accession: A28946
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  A; Molecule type:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                 Query Match
                                                                                                                                                                                                                                                                                                                                       A; Residues:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            A; Accession:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              833860
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                á
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        qq
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    õ
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          Q
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          Db
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       A.Status: preliminary
A.Molecule type: mRNA
A.Cross references: GB:M31106; NID:g161004; PIDN:AAA29888.1; PID:g552241
R.Henkle, K.J.; Davern, K.M.; Wright, M.D.; Ramos, A.J.; Mitchell, G.F.
MOL Blochem. Parasitol. 40, 23-34, 1990
A.Title: Comparison of the cloned genes of the 26- and 28-kilodalton glutathione S-trans
A.Reference number: A44941; MUID:90271935
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        C;Species: Schistosoma mansoni
C;Date: 03-Jun-1993 #sequence_revision 03-Jun-1993 #text_change 18-Jun-1999
C;Accession: A4552; A46341
R;Trottein, F; Kieny, M.P.; Verwaerde, C.; Torpier, G.; Pierce, R.J.; Balloul, J.M.; Sc
Mol. Biochem. Parasitol. 41, 35-44, 1990
A;Title: Molecular cloning and tissue distribution of a 26-kilodalton Schistosoma manson
A;Reference number: A45523; MUID:90348716
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             A)Cross-references: GB:M26914; GB:M26913; NID:g161006; PIDN:AAA29889.1; PID:g161007 C; Superfamily: glutathione transferase C; Keywords: transferase
                                                                                                                                                             ..
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     ;
0
                                                                                                                                                                                                                                                                     65 VKLTQSMAIIRYIADKHNMLGGCPKERAEISMLEGAVLDIRYGVSRIAYSKDFETLKVDF 124
                                                                                                                                                                                                                                                                                             125 LSKLPEMLKMFEDRLCHKTYLNGDHVTHPDFMLYDALDVVLYMDPMCLDAFPKLVCFKKR 184
                                                                                                                                                                                                                                                                                                                                                                               Gaps
                                                                                                                                                                                                         5 PILGYWKIKGLVQPTRLLLEYLEEKYEEHLYERDEGDKWRNKKFELGLEFPNLPYYIDGD 64
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     125 LSKLPEMLKMFEDRLCHKTYLNGDHVTHPDFMLYDALDVVLYMDPMCLDAFPKLVCFKKR 184
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             65 VKLTQSMAIIRYIADKHNMLGGCPKERAEISMLEGAVLDIRYGVSRIAYSKDFETLKVDF 124
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          0; Gaps
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         5 PILGYWKIKGLVQPTRLLLEYLEEKYEEHLYERDEGDKWRNKKFELGLEFPNLPYYIDGD 64
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    0;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    glutathione transferase (EC 2.5.1.18) - fluke (Schistosoma mansoni)
                                                                                                     22.6%; Score 984; DB 2; Length 218;
82.9%; Pred. No. 5.4e-47;
Live 18; Mismatches 19; Indels
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            21.2%; Score 921; DB 2; Length 218; 80.1%; Pred. No. 1.5e-43; tive 14; Mismatches 29; Indels
                 A:Note: sequence extracted from NCBI backbone (NCBIP:77814)
C;Superfamily: glutathione transferase
                                                                                                                                                                                                                                                                                                                                                                                                                                      185 IEAIPQIDKYLKSSKYIAWPLQGWQATFGGGDHPPK 220
                                                                                                                                                                                                                                                                                                                                                                                                                                                               183 IENLPPIKNYLNSNRYIKWPLQGWSATFGGGDAPPK 218
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     IEAIPQIDKYLKSSKYIAWPLQGWQATFGGGDHPPK 220
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           183 IEDLPQIKNYLNSSRYIKWPLQGWDATFGGGDTPPK 218
                                                                                                                                           Matches 179; Conservative
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   Matches 173; Conservative
                                                                                                                         Similarity
A; Residues: 1-218 <WRI>
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             Best Local Similarity
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                A; Residues: 8-218 <HEN>
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          preliminary
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            A; Molecule type: mRNA
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          A; Accession: A45523
                                                                                                   Query Match
                                                                                                                         Best Local
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              Query Match
                                                                                                                                                                                                                                                                     δ
                                                                                                                                                                                                                                                                                                          q
                                                                                                                                                                                                                                                                                                                                                    ò
                                                                                                                                                                                                                                                                                                                                                                                            g
                                                                                                                                                                                                                                                                                                                                                                                                                                        ô
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             ô
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     q
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 õ
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         g
```

```
C;Accession: A4888
R;Muro, A.; Rodriguez-Medina, J.R.; Hillyer, G.V.
Am. J. Trop. Med. Hygy, 48, 477-463, 1993
A. Title: Sequence analysis of a Fasciola hepatica glutathione S-transferase cDNA clon A;Reference number: A48388; MUID:93228188
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                ö
glutathione S-transferase - liver fluke (fragment)
C;Species: Fasciola hepatica (liver fluke)
C;Date: 19-Nov-1993 #sequence_revision 18-Nov-1994 #text_change 03-May-1996
                                                                                                                                                                                                                                                                                                               A:Note: sequence extracted from NCBI backbone (NCBIN:129138, NCBIP:129139)
C:Superfamily: glutathione transferase
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 72 AIIRYIADKHUMLGGCPKERAEISMLEGAVLDIRYGVSRIAYSKDFETLKVDFLSKLPEM 131
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  132 LKMFEDRLCHKTYLNGDHVTHPDFMLYDALDVVLYMDPMCLDAFPKLVCFKKRIBAIPQI 191
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         12 IKGLVQPTRLLLEYLEEKYEEHLYERDEGDKWRNKKFELGLEFPNLPYYIDGDVKLTQSM 71
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    Length 209;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                      56; Indels
                                                                                                                                                                                                                                                                                                                                                                                                                   Score 654; DB 2;
Pred. No. 5.1e-29;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                 31; Mismatches
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        181 KAYMESEKFIKWPLNSWIASFGGGDAAP 208
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               192 DKYLKSSKYIAWPLQGWQATFGGGDHPP 219
```

glittathione transferase (EC 2.5.1.18) mul - mouse
N.Alternate names: glutathione transferase major chain; glutathione transferase mu-8.
C.Species: Mus musculus (house mouse)
C.Date: 08-Dec-1993 #sequence_revision 07-oct-1994 #text_change 18-Jun-1999
C.Accession: S33860; A34159; A28946; A20831; 124735; S16933; S71312
Arch. Biochem. Biophys. 303, 38-393, 1993
A.Title: The structure of two murine class-mu glutathione transferase genes coordinat

A.Residues: 1-218 https://documents.com/residue A.Cross-references: GB:L13448
A.Note: the authors translated the codon GAG for residue 49 as Gly and GAG for residu A.V.; Goldsmith, M.E.; Pickett, C.B.; Cowan, K.H.
J. Biol. Chem. 264, 21582-21590, 1989
A.Title: Isolation, characterization, and expression in Escherichia coli of two murin A.Reference number: A34159; MUID:90094327

A;Molecule type: mRNA A;Residues: 1-218 <TOW> A;Cross-references: GB:J04632; NID:9193547; PIDN:AAA37705.1; PID:9309260

R:Pearson, W.R.; Reinhart, J.; Sisk, S.C.; Anderson, K.S.; Adler, P.N.
J. Blol. Chem. 263, 1334–13332, 1988
A:Fitle: Tissue-specific induction of murine glutathione transferase mRNAs by butylat
A:Reference number: A92668; MUID:88330838

```
A.Wolecule type: Protein
A.Wolecule type: Protein
A.Wolecule type: Protein
A.Wolecule type: Protein
A.Katusz, R.M.; Bono, B.; Collana, R.F.
Arch Blochem. Blophys. 298, 667-677, 1992
A.Title: Identification of Tyr(115) labeled by S-(4-bromo-2,3-dioxobutyl)glutathione
A.Reference number: S27111; MUID:93037509
A.Accession: S27111; MUID:93037509
A.Molecule type: Protein
A.Residues: 2-20;83-86/x', 88-96,109-115,'X',117-122 <KAT>
A.Residues: 2-20;83-86/x', 88-96,109-115,'X',117-122 <KAT>
Blochem J. 278, 293-29', 1991
A.Title: Cysteine-86 is not needed for the enzymic activity of glutathione S-transfer
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                A;Accession: B61004
A;Molecule type: protein
A;Residues: 2-26 c<41-8
B;Mannervik, B.; Alin, P.; Guthenberg, C.; Jensson, H.; Tahir, M.K.; Warholm, M.; Jor
Proc. Natl. Acad. Sci. U.S.A. 82, 7202-7206, 1985
A;Title: Identification of three classes of cytosolic glutathione transferase common
A;Reference number: A24735; MUID:86042634
     A)Cross-references: GB:J02810; NID:g204514; PIDN:AAA41293.1; PID:g204515
R;Lai, H.C.J.; Grove, G.; Tu, C.P.D.
Nucleic Acids Res. 14, 6101-6114, 1986
A)Tille: Cloning and Sequence analysis of a cDNA for a rat liver glutathione S-transf
A)Reference number: A25510; MUID:86312882
A;Accession: A25510
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        A,Title: Identification of rat liver glutathione S-transferase Yb subunits by partial lical isoelectric focusing gel.
A,Reference number: A61004; MUID:91031411
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           A;Residues: 2-20 <WAN>
R;Hsieh, J.C.; Liu, L.F.; Chen, W.L.; Tam, M.F.
Biochem. Biophys. Res. Commun. 162, 1147-1154, 1989
A;Title: Expression of Yb-1 glutathione S-transferase using a baculovirus expression A;Reference number: A33397; MUID:89350924
A;Reference number: A33397
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  121 KVDFLSKLPEMLKMFEDRLCHKTYLNGDHVTHPDFMLYDALDVVLYMDPMCLDAFPKLVC 180
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  61 IDGDVKLJQSMAIIRYIADKHNMLGGCPKERAEISMLEGAVLDIRYGVSRIAYSKDFETL 120
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               6 ILGYWKIKGLVQPTRLLLEYLEEKYEEHLYERDEG-----DKWRNKKFELGLEFPNLPYY 60
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               A; Residues: 1-198, NC', 201-218 <DIN>
A; Residues: 1-198, NC', 201-218 <DIN>
A; Cross-references: GB:M11719; NID:9204502; PIDN:AAA41287.1; PID:9204503
                                                                                                                                                                                                                A; Cross-references: GB:X04229; NID:956337; PIDN:CAA27811.1; PID:956338
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               11.1%; Score 481.5; DB 2; Length 218; 44.4%; Pred. No. 1.4e-19;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               A; Experimental source: clone pGTA/C44
R;Chang, L.H.; Hsieh, J.C.; Chen, W.L.; Tam, M.F.
Electrophoresis 11, 589-593, 1990
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            Mismatches
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       A; Reference number: S17167; MUID:91354218
                                                                                                                                                                                                                                                R. Ding, G.J.F.; Lu, A.Y.H.; Pickett, C.B.
J. Biol. Chem. 260, 13268-13271, 1985
A; Reference number: A24085; MUID:86033768
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       A,Molecule type: protein
A;Residues: 2-18 4822>
C;Superfamily: glutathione transferase
C;Keywords: dimer; transferase
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          38;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          Conservative
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   A; Molecule type: protein
A; Residues: 2-20 <MAN>
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               Local Similarity
es 91; Conserv
                                                                                                                                                                             A; Molecule type: mRNA
A; Residues: 1-218 <LAI>
A;Residues: 1-218 <CHA>
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 A;Status: preliminary
                                                                                                                                                                                                                                                                                                                                                        A; Molecule type: mRNA
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         A;Accession: S17167
                                                                                                                                                                                                                                                                                                                                A; Accession: A24085
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               Query Match
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          Rest Loca
Matches
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    δλ
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  QQ
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       ô
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  Op
  glutathione transferase (EC 2.5.1.18) class mu chain 3 - rat
N.Alternate names: glutathione S-transferase Yb1; ligandin
C;Species: Rattus norvegicus (Norway rat)
C;Date: 28-ANG-1989 #sequence_travision 28-Aug-1989 #text_change 18-Jun-1999
C;Accession: A29794; A25510; A24085; B61004; F24735; A33397; S27111; S17167; A26187
C;Accession: A29794; A25510; A24085; B61004; F24735; A33397; S27111; S17167; A26187
J. Biol. Chem. 262, 11901-11903, 1987
J. Biol. chem. 262, 11901-11903, 1987
A;Atitle: Identification of glutathione S-transferase Yb-1 mRNA as the androgen-repressed A;Accession: A29794; MUID:87308179
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              ;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 121 KVDFLSKLPEMLKMFEDRLCHKTYLNGDHVTHPDFMLYDALDVVLYMDPMCLDAFPKLVC 180
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              61 IDGDVKLTQSMAJIRYIADKHNMLGGCPKERAEISMLEGAVLDIRYGVSRIAYSKDFETL 120
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 64 IDGSHKITQSNAILRYLARKHHLDGETEEERIRADIVENQVMDTRWQLIMLCYNPDFEKQ 123
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               6 ILGYWKIKGLVQPTRLLLEYLEEKYEEHLYERDEG ----DKWRNKKFELGLEFPNLPYY 60
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            4 ILGYWNVRGLTHPIRMLLEYTDSSYDEKRYTMGDAPDFDRSQWLNEKFKLGLDFPNLPYL 63
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           A;Introns: 12/3; 38/1; 59/3; 87/1; 120/3; 152/3; 189/3
C;Superfamily: glutathione transferase
C;Keywords: dimer; transferase
F;2-218/Product: glutathione transferase mul #status experimental <MAT>
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            / Match 11.1%; Score 481.5; DB 2; Length 218; Local Similarity 44.4%; Pred. No. 1.4e-19;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    71; Indels
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    38; Mismatches
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    181 FKKRIEAIPQIDKYLKSSKYIAWPL 205
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         91; Conservative
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          A; Molecule type: mRNA
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              Query Match
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   A; Gene: GSTM1
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            Matches
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            RESULT
```

Indels

71.;

Op

ò

g δλ

δλ qq

```
A;Molecule type: protein
A:Residues: 2-21 <PRI>
C;Superfamily: glutathione transferase
C;Keywords: dimer; transferase
                                                                                                                                                                                                                                                                             Local Similarity
Les 95; Conserv
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       A; Molecule type: DNA
A: Posidues: 1-218 <COM>
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 A;Status: preliminary
A;Molecule type: mRNA
A;Residues: 1-218 <ROS>
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    A;Status: preliminary
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              A; Accession: B47486
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           A; Accession: S36782
                                                                                                                                                                                                                                               Query Match
                                                                                                                                                                                                                                                                                                                   Matches
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  RESULT 11
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        A47486
                                                                                                                                                                                                                                                                                                                                                                                                                                                         q_{\mathrm{Q}}
                                                                                                                                                                                                                                                                                                                                                                                          ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 òγ
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    q
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           qq
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  ô
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  qq
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               glutathione transferase (EC 2.5.1.18) class mu chain 7.8 - rabbit glutathione transferase (EC 2.5.1.18) class mu chain 7.8 - rabbit C.Species: Oryctolagus cuniculus (domestic rabbit) C.Species: Oryctolagus cuniculus (domestic rabbit) C.Species: Oryctolagus cuniculus (domestic rabbit) C.Species: Secondary S10300 C.Species: Secondary C.Species: Seconda
                                                                                                                                                                                                                                                                                                                                                                                          C; Species: Homo sapiens (man)
C; Date: 07-Apr-1994 #sequence_revision 07-Apr-1994 #text_change 18-Jun-1999
C; Accession: 332425; S29337
R; Zhong, S.; Spurr, N.K.; Hayes, J.D.; Wolf, C.R.
Biochem, J. 291, 41-50, 1993
A; Title: Deduced amino acid sequence, gene structure and chromosomal location of a novel
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             A; Residues: 1-218 <ZHO>
A; Cross-references: EMBL:X68677; NID:g31934; PIDN:CAA48637.1; PID:g31935
A; Note: the authors translated the codon AAG for residue 182 as Arg and CCA for residue
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      Ĥ
124 KPEFLKTIPEKMKLYSEFLGKRPWFAGDKVTYVDFLAYDILDQYHIFEPKCLDAFPNLKD 183
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      61 IDGDVKLTQSMAIIRYIADKHNMLGGCPKERAEISMLEGAVLDIRYGVSRIAYSKDFETL 120
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      121 KVDFLSKLPEMLKMFEDRLCHKTYLNGDHVTHPDFMLYDALDVVLYMDPMCLDAFPKLVC 180
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             glutathione transferase (EC 2.5.1.18) class mu, GSTM4 (version 2) - human N:Alternate names: glutathione transferase mu4 \,
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      Gaps
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   6 ILGYWKIKGLVQPTRLLLEYLEEKYEEHLYERD-----EGDKWRNKKFELGLEFPNLPYY 60
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  4 ILGYWDIRGLAHAIRLLLEYIDSSYEEKKYIMGGAPDYDRSOWLNEKFKLGLDFPNLPYL 63
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              2,
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               11.0%; Score 478.5; DB 2; Length 218;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             1 Similarity 44.1%; Pred. No. 2e-19;
94; Conservative 37; Mismatches 77; Indels
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             A,Cross-references: GDB:134191; OMIM:138333
A,Map position: 1p13.3-1p13.3
A,Introns: 12/3; 38/1; 59/3; 87/1; 120/3; 152/3; 189/3
C,Superfamily: glutathione transferase
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        184 FISRFEGLEKISAYMKSSRFLDRPLYTRMAVWG 216
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               181 FKKRIEAIPQIDKYLKSSKYIAWPLQGWQATFG 213
                                                                          181 FKKRIEAIPQIDKYLKSSKYIAWPL 205
                                                                                                                    C; Keywords: dimer; transferase
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                Query Match
Best Local Similarity
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          DNA
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        A; Accession: S32425
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   A; Gene: GDB:GSTM4
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               A; Molecule type:
                                                                                                                                                                                                                                                       6
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 C; Genetics:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               Matches
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           RESULT 10
                                                                                                                                                                                                                                                       RESULT
   a
                                                                       ò
                                                                                                                                          g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             qq
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               ò
```

```
C;Species: Homo sapiens (man)
C;Date: 02-Jun-1995 #sequence_revision 02-Jun-1995 #text_change 18-Jun-1999
C;Accession: A47486; B47486; S336782; I37438; S45685
R;Comstock, K.E.; Johnson, K.J.; Rifenbery, D.; Henner, W.D.
J. Biol. Chem. 268, 16958-16965, 1993
A;Title: Isolation and analysis of the gene and cDNA for a human mu class glutathione
A;Reference number: A47486; MUID:93352467
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     R:Ross, V.L.; Board, P.G.
Biochem, J. 294, 373-380, 193
A:Title: Molecular cloning and heterologous expression of an alternatively spliced hu
A:Reference number: S36782; MUID:93384505
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             R;Comstock, K.E.; Widersten, M.; Hao, X.Y.; Henner, W.D.; Mannervik, B. Arch. Biochem. Biophys. 311, 487-495, 1994
A;Title: A comparison of the enzymatic and physicochemical properties of human glutat A;Reference number: $45684; MUID:94263230
                                                                                                                                                                                                                                         62 DGDVKLTQSMAIIRYIADKHNMLGGCPKERAEISMLEGAVLDIRYGVSRIAYSKDFETLK 121
                                                                                                                                                                                                                                                                             122 VDFLSKLPEMLKMFEDRLCHKTYLNGDHVTHPDFWLYDALDVVLYMDPMCLDAFPKLVCF 181
                                                                                                                                                                                                                                                                                                                                                                                                             7 LGYWKIKGLVQPTRLLLEYLEEKYEEHLYERDEG ----- DKWRNKKFELGLEFPNLPYYI 61
                                                                                                                                                                 glutathione transferase (EC 2.5.1.18) class mu, GSTM4 (version 1) - human
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          A;Status: preliminary
A;Molecule type: mRNA
A;Residues: 1-218 <CCQ2>
A;Cross-references: GB:M96234; NID:g306818; PIDN:AAA57347.1; PID:g306819
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   A;Molecule_type: DNA
A;Residues: 39-120 <RES>
A;Cross-references: EMBL:X56837; NID:931936; PIDN:CAA40167.1; PID:931937
                                                                     ŝ
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          A;Contents: annotation
R;Taylor, J.B.; Oliver, J.; Sherrington, R.; Pemble, S.E.
Biochem. J. 274, 587-593, 1991
A;Title: Structure of human glutathione S-transferase class Mu genes.
A;Reference number: 137438, MUID:91174774
A;Accession: I37438
11.0%; Score 477.5; DB 2; Length 218; 44.6%; Pred. No. 2.3e-19:
                                                                 Indels
                             Pred. No. 2.3e-19;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     A:Status: preliminary; translated from GB/EMBL/DDBJ
                                                           36; Mismatches
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                182 KKRIEAIPQIDKYLKSSKYIAWPLQGWQATFGG 214
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    A; Experimental source: Hela cells
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     A:Cross-references: EMBL:M99422
A:Experimental source: testis
                                                           Conservative
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            A;Cross-references: GB:M96233
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            C;Genetics:
```

```
-----
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             A;Molecule type: protein
A;Residues: 2-29,'K',31 <SUZ>
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             A; Experimental source: brain
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        A; Gene: GDB:GSTM2; GST4
                                                                                                                                                                      A; Molecule type: mRNA
A; Residues: 1-218 <VOR>
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           A; Status: preliminary
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          A; Accession: S14344
                                                                                                                                                A; Accession: A39375
                                                                                                                                                                                                                                                                                                                                                                           A; Accession: A28810
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        A; Accession: S13305
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          Query Match
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       C; Genetics
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       Matches
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         δλ
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    ŏ
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            a
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         δλ
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    Dp
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 QΫ́
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         glutathione transferase (EC 2.5.1.18) mu-2 (validated) - human
N;Alternate names: glutathione S-transferase GST4, muscle; glutathione transferase GST6;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           glutathione transferase (EC 2.5.1.18) b - guinea pig
N;Alternate names: glutathione S-alkyltransferase; glutathione S-aryltransferase; glutath
C;Species: Cavia porcellus (guinea pig)
C;Date: 07-Sep-1990 #sequence_revision 07-Sep-1990 #text_change 18-Jun-1993
C;Accession: JX0095
R;Kamei, K.; Oshino, R.; Hara, S.
B;Comem. 107, 111-117, 1990
A;Title: Anino acid sequence of glutathione S-transferase b from guinea pig liver.
A;Reference number: JX0095; MUID:90236961
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       in
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    A Residues: 1-217 <KAM>
C; Comment: Glutathione transferases are a family of enzymes that are multifunctional C; Superfamily: glutathione transferase
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 Ĥ
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      62 DGDVKLTQSMAIIRYIADKHNMLGGCPKERAEISMLEGAVLDIRYGVSRIAYSKDFETLK 121
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          64 DGTHKLTQSNAILRYIARKHNLCGVTEEETIRMDILENQVMDIRMQLIMLCYSPDFEQKK 123
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         122 VDFLSKLPEMLKMFEDRLCHKTYLNGDHVTHPDFMLYDALDVVLYMDPMCLDAFPKLVCF 181
                                                                                                                                                                                                                                                                                                                                                                                           65 DGAHKITQSNAILCYIARKHNLCGETEEEKIRVDILENQAMDVSNQLARVCYSPDFEKLK 124
                                                                                                                                                                                                                                                                                                                                                                                                                                                  122 VDFLSKLPEMLKMFEDRLCHKTYLNGDHVTHPDFMLYDALDVVLYMDPMCLDAFPKLVCF 181
                                                                                                                                                                                                                                                                                                                                                                                                                                                                            5; Gaps
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           4 LGYWNIRGLTHPIRLILEYTNSGYEEKRYNMGDAPDYDRSQWLNEKFKLGLDFPNLFYLI 63
                                                                                                                                                                                                                                                                                                                                                             62 DGDVKLTQSMAIIRYIADKHNMLGGCPKERAEISMLEGAVLDIRYGVSRIAYSKDFETLK 121
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          7 LGYWKIKGLVQPTRLLLEYLEEKYEEHLYERDEG-----DKWRNKKFELGLEFPNLPYYI 61
                                                                                                                                                                                                                                                                     7 LGYWKIKGLVQPTRLLLEYLEEKYEEHLYERDEG-----DKWRNKKFELGLEFPNLPYYI 61
                                                                                                                                                                                                                                                                                               2
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  Length 217;
                                                                                                                                                                             10.9%; Score 474.5; DB 2; Length 218; 43.9%; Pred. No. 3.4e-19; tive 37; Mismatches 77; Indels 5.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    Indels
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    10.9%; Score 473.5; DB 2; 45.1%; Pred. No. 3.8e-19;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 73;
A;Gene: GDB:GSTW4
A;Cross-references: GDB:134191; OMIM:138333
A;Map position: 1p13.3-1p13.3
A;Introns: 12/3; 38/1; 59/3; 87/1; 120/3; 152/3; 189/3
C;Superfamily: glutathione transferase
C;Superfamily: glutathione splicing; dimer; transferase
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           Query Match
Best Local Similarity 45.18; Pred. ...
Watches 92; Conservative 34; Mismatches
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      | | : : | |:|||:|| | :|| | 185 ISRFEGLEKISAYMKSSRFLPKPLYTRVAVWG 216
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               182 KKRIEAIPQIDKYLKSSKYIAWPLQGWQATFG 213
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 184 ISRFEGLEKISSYMKSSRFLPKPL 207
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      182 KKRIEAIPQIDKYLKSSKYIAWPL 205
                                                                                                                                                                                                                                 93; Conservative
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               A; Molecule type: protein A; Residues: 1-217 < KAM>
                                                                                                                                                                                                         Best Local Similarity
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         C; Keywords: transferase
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         A; Accession: JX0095
                                                                                                                                                                                      Query Match
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   RESULT 13
                                                                                                                                                                                                                                      Matches
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              RESULT
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             A39375
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      δ
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              pp
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   a
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   δλ
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        ŏ
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           δy
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      g
                                                                                                                                                                                                                                                                                                                                                                      δ
                                                                                                                                                                                                                                                                                                                                                                                                                 g
                                                                                                                                                                                                                                                                                                                                                                                                                                                         ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   δλ
                                                                                                                                                                                                                                                                                                                         g
                                                                                                                                                                                                                                                                                 ò
```

```
A.Rolecule type: protein
A.Residues: 2-7, X',9-10, X',12-14 <SIN>
A.Residues: 2-7, X',9-10, X',12-14 <SIN>
A.Residues: 2-7, X',9-10, X',9-10, X',12-14 <SIN>
B.Tsuchida, S., Maki, T.: Sato, K.
J. Biol. Chem. 265, 7150-7157, 1990
A.Title: Purification and characterization of glutathione transferases with an activi A.Reicrence number: A3187; MUID:90237002
A.Roccession: C35187
A.Rolecule type: protein
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   R;Suzuki, T.; Shaw, D.C.; Board, P.G.
Biochem. J. 274, 405-408, 1991
A;Title: Purification and characterization of acidic glutathione S-transferase 6 from
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                R; Hussey, A.J.; Kerr, L.A.; Cronshaw, A.D.; Harrison, D.J.; Hayes, J.D. Biochem. J. 273, 323-332, 1991
A; Title: Variation in the expression of Mu-class glutathione S-transferase isoenzymes A; Reference number: S13305; MUID:91119559
R;Singh, S.V.; Ahmad, H.; Kurosky, A.; Awasthi, Y.C. Arch. Bloochem. Bloochem. 264, 13-22, 1988
A;Tille: Purification and characterization of unique glutathione S-transferases from A;Reference number: A90081; MUID:88280250
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       C;Keywords: brain; cardiac muscle; heart; homodimer; skeletal muscle; transferase F;2-218/Product: glutathione transferase mu-2 #status experimental <MAT>
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      62 DGDVKLTQSMAIIRYIADKHNMLGGCPKERAEISMLEGAVLDIRYGVSRIAYSKDFETLK 121
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   65 DCTHKITQSNAILRYIARKHNLCGESEKEQIREDILENQEMDSRMQLAKLCYDPDFEKLK 124
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             122 VDFLSKLPEMLKMFEDRLCHKTYLNGDHVTHPDFMLYDALDVVLYMDPMCLDAFPKLVCF 181
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 5; Gaps
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            7 LGYWKIKGLVQPTRLLLEYLEEKYEEHLYERDEG-----DKWRNKKFELGLEFPNLPYYI 61
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       5 LGYWNIRGLAHSIRLLLEYTDSSYEEKKYTMGDAPDYDRSQWLNEKFKLGLDFPNLPYLI 64
                                                                                                                                                                                                                                                                                                                                         A; Cross-references: GB:M63509; NID:9183300; PIDN:AAA60963.1; PID:9183301
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     Length 218;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                Indels
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     / Match 10.9%; Score 473.5; DB 2; Local Similarity 44.3%; Pred. No. 3.8e-19; nes 94; Conservative 35; Mismatches 78;
                                                                                                                                                                                                                                                                                                                                                                                      A;Experimental source: myoblast
A;Note: translation of initiator Met is not shown
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   182 KKRIEAIPQIDKYLKSSKYIAWPLQGWQATFG 213
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  A, Cross-references: GDB:128593; OMIM:138380
A, Map position: 1p13:3-1p13.3
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      A; Reference number: S14344; MUID: 91174747
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                A;Molecule type: protein
A;Residues: 2-25 <HUS>
A;Experimental source: skeletal muscle
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    C; Superfamily: glutathione transferase
```

```
121 KVDFLSKLPEMLKMFEDRLCHKTYLNGDHVTHPDFMLYDALDVVLYMDPMCLDAFPKLVC 180
                                                                                                                                                   124 KPEFLKTIPEKMKMYSEFLGKRPWFAGDKVTLCGFLAYDVLDQYQMFEPKCLDPFPNLKD 183
                                                                                                                                                                                                                                                                                                                                                                                    Search completed: January 15, 2002, 13:52:18 Job time: 387 sec
                                                                                                                                                                                                            181 FKKRIEAIPQIDKYLKSSKYIAWPL 205
                                                                                                                                                                                                                                                                    184 FLARFEGLKKISAYMKTSRFLRRPI 208
                                                                                                 οy
                                                                                                                                                   q
                                                                                                                                                                                                            δ
                                                                                                                                                                                                                                                     R;Norris, J.S.; Schwartz, D.A.; MacLeod, S.L.; Fan, W.; O'Brien, T.J.; Harris, S.E.; Tri
Mol. Endocrinol. 5, 979-986, 1991
A;Title: Cloning of a mu-class glutathione S-transferase complementary DNA and character
A;Reference number: A23732; MUID:92049380
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    A;Molecule type: mRNA

A;Residues: 1-218 <ROB>

A;Cross-references: GB:X57489; NID:g49541; PIDN:CAA40726.1; PID:g49542

A;Note: in Genbank entry CLY1, release 109, the source is designated as Cricetulus longi

C;Superfamily: glutathione transferase

C;Keywords: dimer; transferase
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        de
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              Glutathione transferase (EC 2.5.1.18) Y1 - Chinese hamster C:Species: Cricetulus griseus (Chinese hamster)
C:Date: 18-Feb-1994 #sequence_revision 06-Jan-1995 #text_change 18-Jun-1999
C:Accession: S1320
C:Accession: S1320
C:Accession: Jail Wincent, Bar. Hyrien, O.; Debatisse, M.; Buttin, G.
Eur. J. Biochem 193, 19-24, 1990
A:Title: Coamplification of mu class glutathione S-transferase genes and an adenylate
A:Reference number: S13202; MUID:91031445
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            2;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         ..
;-
                                                                                                                                                                     C;Species: Mesocricetus auratus (golden hamšter)
C;Date: 30-Dec-1991 #sequence_revision 30-Dec-1991 #text_change 18-Jun-1999
C;Accession: A23732; S21908
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          PYYIDGDVKLTQSMAIIRYIADKHNMLGGCPKERAEISMLEGAVLDIRYGVSRIAYSKDF 117
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      118 ETLKVDFLSKLPEMLKMFEDRLCHKTYLNGDHVTHPDFMLYDALDVVLYMDPMCLDAFPK 177
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               Gaps
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            Gaps
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   61 IDGDVKLTQSMAIIRYIADKHNMLGGCPKERAEISMLEGAVLDIRYGVSRIAYSKDFETL 120
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            4 LPI - LGYWKIKGLVQPTRLLLEYLEEKYEEHLYERDEG----- DKWRNKKFELGLEFPNL 57
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     6 ILGYWKIKGLVQPTRLLLEYLEEKYEEHLYERDEG-----DKWRNKKFELGLEFPNLPYY 60
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          4 ILGYWNVRGLINPIRLLLEYTDSSYEEKKYTMGDAPDSDRSQWLNEKFKLGLDFPNLPYL 63
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   A)Cross-references: EMBL:X61033; NID:949638; PIDN:CAA43368.1; PID:949639 A;Note: submitted to the EMBL Data Library, July 1991 C;Superfamily: glutathione transferase C;Keywords: transferase
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      5,
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            9
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 Length 218;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 Length 218;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         68; Indels
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      Indels
                                                                                                                                          glutathione transferase (EC 2.5.1.18) mu - golden hamster
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              DB 2;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        Query Match
10.8%; Score 470; DB 2;
Best Local Similarity 44.3%; Pred. No. 5.9e-19;
Matches 90; Conservative 39; Mismatches 68,
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     10.8%; Score 468.5; DB 2 43.4%; Pred. No. 7.2e-19;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                36; Mismatches
185 ISRFEGLEKISAYMKSSRFLPRPVFTKMAVWG 216
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       LVCFKKRIEAIPQIDKYLKSSKY 200
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    Query Match
10.8%
Best Local Similarity 43.4%
Matches 89; Conservative
                                                                                                                                                                                                                                                                                                                                                                                                A;Status: preliminary
A;Molecule type: mRNA
A;Residues: 1-218 <NOR>
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  A; Accession: S13202
A; Status: preliminary
                                                                                                                                                                                                                                                                                                                                                                        A; Accession: A23732
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         58
q
```

g οy qq ò Q ŏ q

g ò

δ

ò

GenCore version 4.5 Copyright (c) 1993 - 2000 Compugen Ltd.	OM protein - protein search, using sw model	January 15, 2002, 14:06:15 ; Search time 134.53 Seconds (Without alignments) 713.258 Million cell updates/sec	US-09-724-296-36 e: 3482 l MPSRKSKAAALDTPQSESSTREVKKGKVPEEVEDEGEFDG 656	e: BLOSUM62 Gapop 10.0 , Gapext 0.5	473505 seqs, 146272329 residues	Total number of hits satisfying chosen parameters: 473505	eq length: 0 eq length: 2000000000	Post-processing: Minimum Match 0% Maximum Match 100% Listing first 45 summaries	<pre>SPTREMBL_17:* 1: sp_archea:* 2: sp_bacteria:* 3: sp_loatceria:* 4: sp_human:* 5: sp_lovercebrate:* 6: sp_nammal:* 7: sp_mhc:* 10: sp_plant:* 10: sp_plant:* 11: sp_virus:* 12: sp_virus:* 13: sp_vertebrate:* 14: sp_unclassified:*</pre>
Copyri	otein sea	January	US-09-7 3482 1 MPSRK	BLOSUM6 Gapop 1	473505	hits sat	ength: 0	Minimum Maximum Listing	#
	OM protein - pro	Run on:	Title: Perfect score: Sequence:	Scoring table:	Searched:	Total number of	Minimum DB seq l Maximum DB seq l	Post-processing:	Database :

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

		æ				
Result No.	Score	Query Match	Query Match Length DB	DB	ID	Description
-	3482	100.0	959	3	001408	On 1408 mental and the second of the second
~'	1011.5	29.0	599	٣	010988	O10988 schizosacch
<u>~</u>	377	10.8	322	7	Q9K9P8	Ogkon8 harillus ha
_	359	10.3	317	7	09L4F0	0914f0 bacillus ce
'n	359	10.3	325	7	O9L4E9	Ogl4P9 bacillus og
G	261.5	7.5	305	7	6M0S60	OgsOme deinococcus
~	261.5	7.5	326	7	OPRIE	Ofrtef deinococcus
æ	160.5	4.6	1280	4	075158	075158 homo sanjen
on.	160.5	4.6	1341	4	09UKV3	Oguky3 homo sanien
_	155.5	4.5	5327	S	076891	076891 drosophila
_	152.5	4.4	1763	11	Q9JKX5	ufuosum sum 60
~	144	4.1	16215	2	O9NFS3	Ognisa drosophila
<u>س</u>	143.5	4.1	828	'n	OSVEUR	Ogven8 drosophila
-	143.5	4.1	914	4	09H945	O9h945 homo sanjen
10	141.5	4.1	1110	13	091255	091255 notromyson
١.	141	4.0	776	4	O9H1R4	Oghlad bomo sanion
_	140.5	4.0	1338	=	8×11'60	Ogity mond sapten
_	140.5	4.0	3263	'n	0917113	Ogi7n3 drocophila
~	140.5	0.4	6815	ď	0017114	Obized dropophila
			1)	*O. T.C.	Carladosomo #n/re

061802 caenorhabdi 09k0b4 neisseria m 01687 homo sapien 026471 schistocerc 026471 schistocerc 090mt5 arabidopsis 09c3y7 candida alb 09c4n4 drosophila 09c3y7 candida alb 09th4 drosophila 09f86 mus musculu 09f856 homo sapien 08f8737 mus musculu 09f856 nemorhabdi 09ex77 drosophila 09ex78 caenorhabdi 09ex78 caenorhabdi 09ex78 caenorhabdi 09ex78 drosophila 09ex78 drosophila 09ex79 drosophila 09ex79 drosophila 09ex70 drosophila 09ex70 mus musculu 09ex70 musculu	
061802 09K0B4 09K0B4 09UKXO 09UKXO 09VKBB 09C3Y7 099C682 09C3Y7 099C682 09C3Y1 09F0B6	
11510 5 11815 2 11815 2 11781 4 11547 5 806 1 1137 3 11591 1 11591 1 11591 5 11791 5 1	
44444000000000000000000000000000000000	
139.5 138.5 138.5 138.5 137.5 137.5 137.5 137.5 137.5 137.5 137.5 138.5 13	
011242444444444444444444444444444444444	

ALIGNMENTS

 $^{\sim}$

us-09-724-296-36.rspt

```
ENDDVGWTVHDLLPVCEELNIPMVLDYHHHNICFDPAHLREGTLDISDPKLQERIANTWK 480
                                                                                                                                                                                                                                                              RKGIKQKMHYSEPCDGAVTPRDRRKHRPRVMTLPPCPPDMDLMIEAKDKEQAVFELMRTF 540
                                                                                                                                                                                                                                                                                                                                                                     541 KLPGFEKINDMVPYDRDDENRPAPPVKAPKKKKGGKRKRTTDEEAAEPEEVDTAADDVKD 600
                 241 PLQFEDEPEHHLKNKPDKSKEPQDELGHKFVQELGLANARDIVKMLCWNEKYGIRFLRLS 300
                                                                                                                        SEMFPFASHPVHGYKLAPFASEVLAEAGRVAAELGHRLTTHPGQFTQLGSPRKEVVESAI 360
                                                                                                                                          361 RDLEYHDELLSLLKLPEQQNRDAVMIIHMGGQFGDKAATLERFKRNYARLSQSCKNRLVL 420
181 AQDAAERGAARPPAVNSSYLPLPWKGRLGYACLNTYLRNAKPPIFSSRTCRMASIVDHRH 240
                                                                                         241 PLQFEDEPEHHLKNKPDKSKEPQDELGHKFVQELGLANARDIVKMLCWNEKYGIRFLRLS 300
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           Takao M., Yonemasu R., Yamamoto K., Yasui A.;
Ticharacterization of a UV endonuclease gene from the fission yeast
Schizosaccharomyces pombe and its bacterial homolog.";
Nucleic Acids Res. 24:1267-1271(1996).
                                                                                                                                                                                                                                                                                                                           601 APEGPKEVPEEERAMGGPYNRVYWPLGCEEWLKPKKREVKKGKVPEEVEDEGEFDG 656
                                                                                                                                                                                                                                                                                                                                                                                                                                                  Lyne M., Rajandream M.A., Barrell B.G., Lucas M., Gaillardin C., Submitted (JUN-1998) to the EMBL/GenBank/DDBJ databases.
EMBL; D78571; BAA11415.1; -.
EMBL; U78487; AAC49664.1; -.
                                                                              Davey S., Nass M.L., Ferrer J.V., Sidik K., Eisenberger A., Mitchell D.L., Freyer G.A.;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 "The fission yeast UVDR DNA repair pathway is inducible."; Nucleic Acids Res. 25:1002-1008(1997).
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         599 AA; 68816 MW; C09ED7219D5895BE CRC64;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    Schizosaccharomyces pombe (Fission yeast).
Eukaryota; Fungi; Ascomycota; Schizosaccharomycetes;
Schizosaccharomycetales; Schizosaccharomycetaceae;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         (TrEMBLrel. 01, Last sequence update) (TrEMBLrel. 08, Last annotation update)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               599 AA
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          01-NOV-1996 (TrEMBLrel. 01, Created)
01-NOV-1996 (TrEMBLrel. 01, Last sequ
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               PRT;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        MEDLINE=97175806; PubMed=9023111;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              MEDLINE=96188860; PubMed=8614629;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              EMBL; AL023859; CAA19577.1;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               PRELIMINARY;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      Schizosaccharomyces.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    SEOUENCE FROM N.A.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 SEQUENCE FROM N.A.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           SEQUENCE FROM N.A.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           UV-ENDONUCLEASE.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    NCBI_TaxID=4896;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             Q10988; P87339;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          UVDE OR UVE1+.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    STRAIN=972H-;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              Endonuclease
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 STRAIN-SP972
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           01-NOV-1998
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             SEQUENCE
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               7
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               010988
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 ~
                                                                                                                                                                                                                                                 421
                                                                                                                            301
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                RESULT
Q10988
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                g
                                                                                                                                                                                     ò
                                                                                                                                                                                                                    g
                                                                                                                                                                                                                                               ò
                                                                                                                                                                                                                                                                             g
                                                                                                                                                                                                                                                                                                             δy
                                                                                                                                                                                                                                                                                                                                        P
                                                                                                                                                                                                                                                                                                                                                                        δ
                                                                                                                                                                                                                                                                                                                                                                                                      a
                                                                                                                                                                                                                                                                                                                                                                                                                                     ŏλ
                                                                                                                                                                                                                                                                                                                                                                                                                                                                pp
                                                              δ
                                                                                                                            ŏ
```

```
18;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              427 VKARLVLENDDVSWSVQDLLPLCQELN1PLVLDWHHHNIV--PGTLREGSLDLM--PLIP 482
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  483 TIRETWTRKGITQKQHYSESADPTAISGMKRRAHSDRVFDFPPCDPTMDLMIEAKEKEQA 542
                                                                                                                                                                                                                                                    115 GIPVPSVVSRF-PTAPYHHKSTNAEEREAKEPVLKTHSKDVEREAEIGVDDVVKMEPAAT 173
                                                                                                                                                                                                                                                                                                                                                                       233 ----PREMFDCLDK-------PIPWRGRLGYACLNTILRSMKERVFCSRTCRIT 275
                                                                                                                                                                                                                                                                                                                                                                                                                 234 SIVDHRHPLQFEDEPEHHLKNKPDKSKEPQDELGHKFVQELGLANARDIVKMLCWNEKYG 293
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  294 IRFLRLSSEMFPFASHPVHGYKLAPFASEVLAEAGRVAAELGHRLTTHPGQFTQLGSPRK 353
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    354 EVVESAIRDLEYHDELLSLLKLPEQQNRDAVMIIHMGGQFGDKAATLERFKRNYARLSQS 413
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               414 CKNRLVLENDDVGWTVHDLLPVCEELNIPMVLDYHHHNICFDPAHLREGTLDISDPKLQE 473
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       474 RIANTWKRKGIKQKMHYSEPCD-GAVTPRDRRKHRPRVMTLPPCPPDMDLMIEAKDKEQA 532
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         533 VFELMRTFKLPGFEKINDMVPYDRDDENRPAPPVKAPKKKKGGKRKRTTDEEAAEPEEVD 592
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 543 VFELCRRYEL-----LEIMGPE--- 565
                                                                                                                                                                   63 AGRMMGKDANGHCLREGKEQEEGVKMAIEGLARMERRLQRATKRQKKQLEED------ 114
                                                                                                                                                                                                                                                                                                                               174 NIIEPEDAQDAAERGAARPPAVNSSYLPLPWKGRLGYACLNTYLRNAKPPIFSSRTCRMA 233
                                                                                                                                                                                                                                                                                                                                                                                                                                                           276 TI------DRDGLESVKQLGTQNVLDLIKLVEWNHNFG 307
                                                                                                                         110 ARSSKKKATDSVSDKIDESVASYDSST-----HLRRSSR-----SKKPVNYNSSSESE- 157
                                                                                                                                                                                                          158 ----SEQISKATKKVKQKEEEEYVEEVDEK 184
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          593 TAADDVKDAPEGPKEVPEEERAMGGPYNRVYWPLGCEEWLKPKKREVKKGKVPEEVEDE 651
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 "Complete genome sequence of the alkaliphilic bacterium Bacillus halodurans and genomic sequence comparison with Bacillus subtills.",
                                       Gaps
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   3 SRKSKAAALDTPQSESSTFSSTLDSSAPSPARNLRRSGRNILQPSSEKDRDHEKRSGEEL 62
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               308 IHFMRVSSDLFPFASHAKYGYTL-EFAQSHLEEVGKLANKYNHRLTMHPGQYTQIASPRE
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        Takami H., Nakasone K., Takaki Y., Maeno G., Sasaki R., Masui N.,
Fuji F., Hirama C., Nakamura Y., Ogasawara N., Kuhara S.,
                                                                                                                                                                                                                                                                                          185 SLKNESSSDEFEPVVPEQLETPISKRRS-----RSSAKNLEKESTMNLDDHA-----
                                       149; Indels 181;
DB 3; Length 599;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              Last annotation update)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             Bacteria; Firmicutes; Bacillus/Clostridium group; Bacillus/Staphylococcus group; Bacillus.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   Last sequence update)
  29.0%; Score 1011.5; DB 36.6%; Pred. No. 2.5e-60; Live 88; Mismatches 149
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            322 AA
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  01-0CT_2000 (TrEMBLrel. 15, Created)
01-0CT-2000 (TrEMBLrel. 15, Last sequ
01-0CT-2000 (TrEMBLrel. 15, Last anno
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            PRT;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  STRAIN=C-125 / JCM 9153;
MEDLINE=20512582; PubMed=11058132;
                                            Matches 241; Conservative
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            PRELIMINARY;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           Bacillus halodurans.
                        Similarity
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       566 --YDQTRDG----
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 SEQUENCE FROM N.A
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     NCBI_TaxID=86665;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 BH2597 PROTEIN.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      Horikoshi K.;
    Query Match
                            Best Local
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    09K9P8;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          Q9K9P8
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       m
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       RESULT
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          09K9P8
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  q
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      qq
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    q
                                                                                                                                                                       ΟY
                                                                                                                                                                                                               qq
                                                                                                                                                                                                                                                                                              Ωp
                                                                                                                                                                                                                                                                                                                                        Qγ
                                                                                                                                                                                                                                                                                                                                                                                qq
                                                                                                                                                                                                                                                                                                                                                                                                                       ŏ
                                                                                                                                                                                                                                                                                                                                                                                                                                                               q
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         ŏ
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               q
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           ŏ
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  qq
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            οy
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                δŽ
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 Öλ
```

```
01-OCT-2000
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 IN 549
                                                                                                                                                                                      548 IN 549
                                                                                                                                                                                                    |:
310 IS 311
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 SEQUENCE
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   Query Match
                                                                                                                                                                                                                                                                                  Q9L4E9;
                                                                                                                                                                                                                                                                     Q9L4E9
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           Matches
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  548
                                                                                                                                                                                                                                               RESULT
Q9L4E9
                       qq
                                               Óγ
                                                                      a
                                                                                             ò
                                                                                                                QQ
                                                                                                                                          \delta
                                                                                                                                                               q
                                                                                                                                                                                      ŏ
                                                                                                                                                                                                           q
                                                                                                                                                                                                                                                                                  ó
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       qq
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              δŏ
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  QQ
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         QΥ
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                qq
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       Óγ
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            qq
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   δλ
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          qq
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  δλ
                                                                                             10;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   Okstad O., Gominet M., Purnelle B., Rose M., Lereclus D., Kolsto A.B.; "Sequence analygis of three Bacillus cereus loci under PicR-regulated genes encoding degradative enzymes and enterotoxin."; Microbiology 145:3129-3138 (1999).

EMBL: AJ243712; CAB69813.1; -. Endourclease.

Endourclease. 317 Aa; 36990 MW; A6AC35F5800A22C3 CRC64;
                                                                                                                                                                                                                                                                                                                                                                           307
                                                                                                                207 RLGYACLNTYLRNAKPPIFSSRTCRMASIVDHRHPLQFEDEPEHHLKNKPDKSKEPQDEL 266
                                                                                                                                                              267 GHKFVQELGLANARDIVKMLCWNEKYGIRFLRLSSEMFPFASHPV-HGYKLAPFASEVLA 325
                                                                                                                                                                                                           326 EAGRVAAELGHRLTTHPGQFTQLGSPRKEVVESAIRDLEYHDELLSLLKLPEQQNRDAVM 385
                                                                                                                                                                                                                        501
                                                                                                                                                                                                                                                                                                                              263
                                                                                                                                                                                                                                                                                                                                                    554
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        260 KEPQDELGHKFVQELGLANARDIVKMLCWNEKYGIRFLRLSSEMFPFASHPVHGYKLAPF 319
                                                                                              Gaps
                                                                                                                                       386 IIHMGGQFGDKAATLERFKRNYARLSQSCKNRLVLENDDVGWTVHDLLPVCEELNIPMVL
                                                                                                                                                                                                                                                                                                    DYHHHNICFDPAHLREGTLDISDPKLQERIANTWKRKGIKQKMHYSEPCDGAVTPR----
                                                                                                                                                                                                                                                                                                                     213 DIHHHDVLHRSKSLQE-----TWQRIVATWEDSPLPVKIHLSSPLSGEDDPRHDY
                                                                                                                                                                                                                                                                                                                                                                           -----DLAEY
                                                                                                                                                                                                                                                                                                                                                 502 --- DRRKHRPRVMTLPPCPPD----MDLMIEAKDKEQAVFELMRTFKLPGFEKINDMVPY
                                                                                             68;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   36;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          Query Match
10.3%; Score 359; DB 2; Length 317;
Best Local Similarity 28.8%; Pred. No. 1.3e-16;
Matches 87; Conservative 60; Mismatches 119; Indels
                                                                     Length 322;
                                                                                            Indels
                                   322 AA; 36900 MW; 4B0DC9291CD1E7B4 CRC64;
                                                                                                                                                                                                                                                                                                                                                               Last sequence update)
Last annotation update)
                                                                     Query Match 10.8%; Score 377; DB 2; Le
Best Local Similarity 29.1%; Pred. No. 7.8e-18;
Matches 105; Conservative 61; Mismatches 127;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        Bacteria; Firmicutes; Bacillus/Clostridium group;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                              317 AA
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  Bacillus/Staphylococcus group; Bacillus
Nucleic Acids Res. 28:4317-4331(2000)
EMBL; AP001516; BAB06316.1; -.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     Created)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  SEQUENCE FROM N.A.
STRAIN-ATCC 14579 TYPE STRAIN;
MEDLINE-20055637; Pubmed-10589720;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 01-OCT-2000 (TrEMBLrel. 15, 01-OCT-2000 (TrEMBLrel. 15, 01-MAR-2001 (TrEMBLrel. 16,
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      PUTATIVE UV-ENDONUCLEASE.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                              PRELIMINARY;
                       Complete proteome.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             NCBI_TaxID=1396;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           Bacillus cereus
                                                                                                                                                                                                                                                                                                                                                                                              555 D 555
                                                                                                                                                                                                                                                                                                                                                                                                                     308 D 308
                                 SEQUENCE
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        Q9L4F0;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                              Q9L4F0
                                                                                                                                                                                                                                                                                                                                                                                                                                                        4
                                                                                                                                                                                                                                                                                                      446
                                                                                                                                                                                                                                                                                                                                                                                                                                                                  RESULT
  RE
SO WE
                                                                                                                 ò
                                                                                                                                                              ò
                                                                                                                                                                                   q
                                                                                                                                                                                                           ò
                                                                                                                                                                                                                                 8
                                                                                                                                                                                                                                                        ò
                                                                                                                                                                                                                                                                            a
                                                                                                                                                                                                                                                                                                    ò
                                                                                                                                                                                                                                                                                                                         g
                                                                                                                                                                                                                                                                                                                                                ò
                                                                                                                                                                                                                                                                                                                                                                       엽
                                                                                                                                                                                                                                                                                                                                                                                              ò
                                                                                                                                                                                                                                                                                                                                                                                                                   g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                g
```

```
SEQUENCE FROM N.A.
STRAIN-ATCC 14579 TYPE STRAIN;
MEDLINE-20055637; PubMed=10589720;
MEDLINE-20055637; PubMed=10589720;
MEDLINE-20055637; PubMed=10589720;
Sequence analysis of three Bacillus cereus loci under PICR-regulated genes encoding degradative enzymes and enterotoxin.";
Microbiology 145:3129-3138 (1999).
EMBL: AJ243712; CAB69814.1; -.
320 ASEVLAEAGRVAAELGHRLTTHPGQFTQLGSPRKEVVESA1RDLEYHDELLSLLKLPEQQ 379
                                                                          89 FTPLWRKIGALIKEHNLRISFHPNQFTLFTSDKPHITTNAITDMTYHYKILDAIGIAD-- 146
                                                                                                                                                                                                                                                                           380 NRDAVMIIHMGGQFGDKAATLERFKRNYARLSQSCKNRLVLENDDVGWTVHDLLPVCEEL 439
                                                                                                                                                                                                                                                                                                                                                             | : | | : | : | : | : | : | : | | : | | : | | : | | : | | : | | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : 
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    205 NIPFVFDYHHHMANLCEEPLE------249
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              250 RSEKEFRAHAEYIDLEFIKPFLHIAKKHNHNFDIMIESKQKDLALFQLIDELSAIRGIKR 309
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          380 NRDAVMITHMGGQFGDKAATLERFKRNYARLSQSCKNRLVLENDDVGWTVHDLLPVCEEL 439
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      440 NIPMVLDYHHH--NICFDPAHLREGTLDISDPKLQERIANTWKRKGIKQKMHYSEPCDGA 497
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 498 VTPRDRRKH------RPRVMTLPPCPPDMDLMIEAKDKEQAVFELMRTFK-LPGFEK 547
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            260 KEPQDELGHKFVQELGLANARDIVKMLCWNEKYGIRFLRLSSEMFPFASHPVHGYKLAPF 319
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               320 ASEVLAEAGRVAAELGHRLTTHPGQFTQLGSPRKEVVESAIRDLEYHDELLSLLKLPEQQ 379
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  155 --SSYINIHVGGAYGNKEKAIERFHENIKKLPTHIKKQMTLENDDKTYTTAETLSICQKE 212
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     440 NIPMVLDYHHH--NICFDPAHLREGTLDISDPKLQERIANTWKRKGIKQKMHYSEPCDGA 497
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        498 VIPRDRRKH------RPRVMTLPPCPPDMDLMIEAKDKEQAVFELMRTFK-LPGFEK 547
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  258 RSEKEFRAHAEYIDLEFIKPFLHIAKKHNHNFDIMIESKQKDLALFQLIDELSAIRGIKR 317
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           213 NIPFVEDYHHHMANLCEEPLE-----ELLPAIFETWSHTNISPKVHISSP---- 257
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      36; Gaps
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             10.3%; Score 359; DB 2; Length 325; 28.8%; Pred. No. 1.3e-16; ative 60; Mismatches 119; Indels
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        325 AA; 37900 MW; 07EFBDD1FA580653 CRC64;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               Last sequence update)
Last annotation update)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     Bacteria; Firmicutes; Bacillus/Clostridium group;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               325 AA
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  Bacillus/Staphylococcus group; Bacillus.
NCBI_TaxID=1396;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            Created)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      PRT;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      01-OCT-2000 (TrEMBLrel, 15, 01-OCT-2000 (TrEMBLrel, 15, 01-MAR-2001 (TrEMBLrel, 16,
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              87; Conservative
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             PUTATIVE UV-ENDONUCLEASE.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       PRELIMINARY;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            Local Similarity
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           Bacillus cereus
```

```
Ketchum K.A., Nelson K.E., Salzberg S., Smith H.O., Venter J.C.,
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          TISSUE=BRAIN;
            Fraser C.M.;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     SEQUENCE
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       InterPro;
                                                                                                    SEQUENCE
                                                                                                                                          Query Match
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     KIAA0670.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        075158
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         NON_TER
                                                                                                                                                                                                                                                                                                                                                                                                                                                                  RESULT
                                                                                                                                                                                                                                                                                                                                                                                                                                                                             075158
 RA
RA
RT
RE
DR
DR
SO
                                                                                                                                                                                                                                                                                               δ
                                                                                                                                                                                                                                                                                                                       qq
                                                                                                                                                                                                                                                                                                                                               οy
                                                                                                                                                                                                                                                                                                                                                                         QQ
                                                                                                                                                                                                                                                                                                                                                                                                    Qγ
                                                                                                                                                                                                                                                                                                                                                                                                                           qq
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       13;
                                                                                                                                                                                                                               Kitayama S., Kikuchi M., Funayama T., Narumi I., Watanabe H.; "Cloning of structural gene of an alternative incision enzyme for DNA damage in Deinococcus radiodurans.";
                                                                                                                                                                                                                                                                                                                                                                                                    292 YGIRFLRLSSEMFP---FASHPVHGYKLAPFASEVLAEAGRVAAELGHRLTTHPGQFTQL 348
                                                                                                                                                                                                                                                                                                                                                                                                                  124 NSDRPEVRESSVRAMSAHARVMDGLGLARTPWN-----LLLLHGGKGCRGAELAALI--- 175
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        464 LDISDPKLQE---RIANTWKRKGIKQKMHYSEPCDGAVTPRDRRKHRPRVMTLPPCPPDM 520
                                                                                                                                                                                                                                                                                                                                                                                                                                                     349 GSPRKEVVESAIRDLEYHDEL---LSLLKLPEQQNRDAVMIIH--MGGQFGDKAATLERF 403
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       404 KRNYARLSQSCKNRLVLENDDVGWTVHDLLPVCEELNIPMVLDYHHHNICFDPAHLREGT 463
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    176 ----PDLPDPVRLRLGLENDERAYSPAELLPICEATGTPLVFDAHHHVV------HDKL 224
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                225 PDGEDPSVREWVLRARATWOPPE-WOVVHLSNGIEG---PODRR-HSHLIADFPSAYADV 279
                                                                                                                                                                                                                                                                                                                                                                          39; Gaps
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            MEDLINE-20036896; PubMed=10567266; White O., Eisen J.A., Heidelberg J.F., Hickey E.K., Peterson J.D., Dodson R.J., Haft D.H., Gwinn M.L., Nelson W.C., Richardson D.L., Moffat K.S., Qin H., Jiang L., Pamphile W., Crosby M., Shen M., Vamathevan J.J., Lam P., McDonald L., Utterback T., Zalewski C., Makarova K.S., Aravind L., Daly M.J., Minton K.W., Fleischmann R.D.,
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     Bacteria; Thermus/Deinococcus group; Deinococcales; Deinococcus.
                                                                                                                                                                   Bacteria; Thermus/Deinococcus group; Deinococcales; Deinococcus
                                                                                                                                                                                                                                                                                                                                               Query Match
7.5%; Score 261.5; DB 2; Length 305;
Best Local Similarity 32.2%; Pred. No. 4.6e-10;
Matches 85; Conservative 38; Mismatches 102; Indels 39
                                                                                                                                                                                                                                                                     to the EMBL/GenBank/DDBJ databases
                                                                                                                                                                                                                                                                                                           305 AA; 33592 MW; B94D333243E2FEA4 CRC64;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       Last sequence update)
Last annotation update)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                326 AA.
                                                                                         Created)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                PRT;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          521 D-LMIEAKDKEQAV--FELMRTFK 541
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 280 PWIEVEAKGKEEAIAALRLMAPFK 303
                                                                                                                                                                                                                                                                     Submitted (OCT-1999) to the EMEL; AB033747; BAA85759.1; -.
                                                                                     01-MAY-2000 (TrEMBLrel. 13, 01-MAY-2000 (TrEMBLrel. 13, 01-MAY-2000 (TrEMBLrel. 13,
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                PRELIMINARY;
                                                               PRELIMINARY;
                                                                                                                                                       Deinococcus radiodurans.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        Deinococcus radiodurans.
                                                                                                                                                                                                       SEQUENCE FROM N.A.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         SEQUENCE FROM N.A.
                                                                                                                            UV-ENDONUCLEASE.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 NCBI_TaxID=1299;
                                                                                                                                                                           NCBI_TaxID=1299;
                                                                                                                                                                                                                                                                                                Endonuclease
318 IS 319
                                                                                                                                                                                                                                                                                                            SEQUENCE
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              O9RTE6;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                Q9RTE6
                                                              09S0M9
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       RESULT 7
                                     RESULT
Q9S0M9
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    Q9RTE6
g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                             g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           Q
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      Qy
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       οy
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   ŏ
                                                                                                                                                                                                                                                                                                                                                                                                                                                     δ
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       δ
```

```
Gaps 13;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        Query Match 4.6%; Score 160.5; DB 4; Length 1280;
Best Local Similarity 19.6%; Pred. No. 0.02;
Matches 153; Conservative 104; Mismatches 315; Indels 209; Gaps
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  246 PDGEDPSVREWVLRARATWQPPE-WQVVHLSNGIEG---PQDRR-HSHLIADFPSAYADV 300
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             464 LDISDPKLQE---RIANTWKRKGIKQKMHYSEPCDGAVTPRDRRKHRPRVMTLPPCPPDM 520
                                                                                                                                                                                                                                                                                                                                                                                                                                                 292 YGIRFLRLSSEMFP---FASHPVHGYKLAPFASEVLAEAGRVAAELGHRLTTHPGQFTQL 348
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     349 GSPRKEVVESAIRDLEYHDEL---LSLLKLPEQQNRDAVMIIH--MGGQFGDKAATLERF 403
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     145 NSDRPEVRESSVRAMSAHARVMDGLGLARTPWN-----LLLLHGGKGGRGABLAALI--- 196
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      404 KRNYARLSQSCKNRLVLENDDVGWTVHDLLPVCEELNIPMVLDYHHHNICFDPAHLREGT 463
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              Ishikawa K., Nagase T., Suyama M., Miyajima N., Tanaka A., Kotani H., Nomura N., Ohara O.; "Prediction of the coding sequences of unidentified human genes. X.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 86 HDIRLYRLSSSLFPMLDLAGDDTGAAVLTHLAPQLL-EAGHAFTDAGVRLLMHPEQFIVL 144
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  100 new cDNA clones from brain which can
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            Homo sapiens (Human).
Bakaryota: Metazooa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammania; Eutheria; Primates; Catarrhini; Hominidae; Homo.
                                                                                                                                                                                                                                                                                                                                                                             39;
                                                                                                                                                                                                                                                                                                     7.5%; Score 261.5; DB 2; Length 326; 32.2%; Pred. No. 5e-10; tive 38; Mismatches 102; Indels 39
"Genome sequence of the radioresistant bacterium Deinococcus
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           145145 MW; 726A40C168D6907B CRC64;
                                                                                                                                                                                                           326 AA; 35693 MW; C4EA0D0AD2C38988 CRC64;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    01-NOV-1998 (TrEMBLrel. 08, Last sequence update) 01-JUN-2001 (TrEMBLrel. 17, Last annotation update)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       PRT; 1280 AA
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        01-NOV-1998 (TrEMBLrel. 08, Created)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               code for large proteins in vitro.";
DNA Res. 5:169-176(1998).
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  521 D-LMIEAKDKEQAV--FELMRTFK 541
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              MFDLINE=98403880; PubMed=9734811;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          : :||| ||:|:
301 PWIEVEAKGKEEAIAALRLMAPFK 324
                                                                                                                                    TIGR; DR1819; ..
Endonuclease; Complete proteome.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   EMBL; AB014570; BAA31645.1; -.
InterPro; IPR000504; RRM.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        KIAA0670 PROTEIN (FRAGMENT).
                                                                Science 286:1571-1577(1999).
EMBL: AE002022: AAF11370.1;
                                                                                                                                                                                                                                                                                                                                            Best Local Similarity 32.2% Matches 85; Conservative
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  The complete sequences of
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          IPR003034; SAP.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       PRELIMINARY;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       Pfam; PF02037; SAP; 1.
SMART; SM00360; RRM; 1.
SMART; SM00513; SAP; 1.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           1280 AA;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     SEQUENCE FROM N.A.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             NCBI_TaxID=9606;
                                       radiodurans Rl.'
```

Qy Dp	1 MPSRKSKAAALDTPQSESSTFSSTLDSSAPSPARNLRRSGRNILQPSSEKDRDH 54 : : : :
Qy Db	55 EKRSGEELAGRAMAGKDANGHCLR
Qy Dp	89 AIEGLARMERRLORATKROKKOLEEDGIPVPSVVSRFPTAPYHHKSTNAEERE- 141
<i>₹</i> 6	142AKEPVLKTHSKDVEREAEIGVDDVVKMEPAATNIIEPED 180 1
o o	
Oy Op	223DIFSSRTCRMASIVDHRHPLQFEDEPEHHLKNKPDKSKEPQDELGHKFVQEL 274
Oy Dp	275 GLANARDIVKMLCWNEKYGIRFLRLSSEMFPFASHPVHGYKLAPFASEVLAEAGRVAAEL 334
Qy Dp	335 GHRLTTH
Qy	377 EQONRDAVMIIHMGGOFGDKAATLERFKRNYARLSQSCKNRLVLENDDVGWTVHDL 432
oy G	433 LPVCEELNIPMVLDYHHHNICFDPAHLREGTLDISDPKLQERIANTWRRKGIKQKMHYSE 492 1 :
Oy Dp	493 PCDGAVTPRORRKHRPRVMTLPPCPPDMDLMIEAKDKEQAVFELMRTFKLPGFEKIND 550
Oy Dp	551 MVPYDRDDENRPAPPKKARGGKRKRTTDEEAAEPEEVDTAADDVKDA 601 111 :
Qy	602 PEGPKEVPEEERAMGGPYNRVYW-PLGCEEWLKPKKREVKKGKVPEEVEDEG 652
δ	E 653
QQ	1203 E 1203
RESULT OPEN OF THE OFFI DE ACCOUNTS OF THE OFFI DE ACCOUNTS OFFI DE ACCOUN	UT 9 V3 O9UKV3 PRELIMINARY; PRT; 1341 AA. O9UKV3; O1-MAY-2000 (TrEMBLrel. 13, Created) O1-MAY-2000 (TrEMBLrel. 13, Last sequence update) O1-JUN-2001 (TrEMBLrel. 17, Last annotation update) O1-JUN-2001 (TrEMBLrel. 17, Last annotation update) ACINUSE. RUKATYOCA; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo. NCBI_TAXID=9606; Ill SEQUENCE FROM N. MEDLINE=99418558; PubMed=10490026; SAHARA S., AOLO M., EQUChi Y., Imamoto N., Yoneda Y., Tsulimoto Y.;

```
1030 ELLGRIGTLV------EEAFWIDKIKSHCFVTYSTVEEAVATRTALHG--VKWPOSNP 1079
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      974 GDTLTRRSISQQKSGVSITIDDPVRTAQVPSPPR----GKISNIVHISNLVRPFTLGQLK 1029
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               1080 KFLCADYAEQDELDYHRGILVDRP-----SETKTEE-----OGIPRPLHPPP 1121
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        433 LPVCEELNIPMVLDYHHHNICFDPAHLREGTLDISDPKLQERIANTWKRKGIKQKMHYSE 492
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       PCDGAVIPRDRRHHRPRVMTLPPCPPDMDLMIEAKDKEQAVFE--LMRTFKLPGFEKIND 550
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    602 PEGPKEVPEEERAMGGPYNRVYW-PL-----GCEEWLKPKKREVKKGKVPEEVEDEG 652
                                                                                                                                                                                                                                                                                                              142 --AKEP----VLKTHSKD-----VEREAEIGVD------DVVKMEPAATNIIEPED 180
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          223 ----PIFSSRTCRMASIVDHRHPLQFEDEPEHHLKNKPDKSKEPQDEL---GHKFVQEL 274
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    857 TTQKKPSISITTESLKSLIPDIKPLAGO---EAVVDLHADDSRISEDETERNGDGTHDK 913
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    275 GLANARDIVKMLCWNEKYGIRFLRLSSEMFPFASHPVHGYKLAPFASEVLAEAGRVAAEL 334
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   914 GLKICRTVTQVVPAEGQENGQREEEEEKEPEAEPPVPPQVSVEVALPPPAEHEVKKVTL 973
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        335 GHRLTTH-----SLLKLP 376
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             377 EQONRDAVMIIHMGGQFGDKAATLERFKRN----YARLSQSCKNRLVLENDDVGWTVHDL 432
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             551 MVPYDRD-----DENRPAPPVKAPKKKKGKRKRTTDEEAAEPEEVDTAADDV----KDA 601
                                                                                                                                                                                                                Ouery Match 4.6%; Score 160.5; DB 4; Length 1341; Best Local Similarity 19.6%; Pred. No. 0.021; Matches 153; Conservative 104; Mismatches 315; Indels 209; Gaps
                                                                                                                                                                                                                                                                                                                                                                                                                               89 AIEGLARM-----ERRLQRATKRQKKQLEEDGIPVPSVVSRFPTAPYHHKSTNAEERE- 141
                                                                                                                                                                                                                                                                                      1 MPSRKSKAAALDTPQSESSTFSSTLDSSAPSPARNLRRSGRNILQPSSEKDRD-----H 54
                                                                                                                                                                                                                                                                                                                                                             55 EKRSGEELAG------RMMGKDANGHCLR--------EGKEQEEGVKM 88
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          181 AQDAAERGAARPPAVNSSYLPLPWKGRL-----GYACLNTYLRNAKP-----
"Acinus is a caspase-3-activated protein required for apoptotic chromatin condensation.";
Nature 401:168-173(1999).
BMBL: AF124726. AAD56724-1; -.
InterPro; IPR00504; RRM.
InterPro; IPR007034; SAP.
Pfam; PF02037; SAP; 1.
SMART; SM00360, RRM; 1.
SMART; SM00513; SAP; 1.
SMART; SM00513; SAP; 1.
SMART; SM00513; SAP; 1.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     E 1264
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     ----d
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          653 E 653
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       493
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        1122
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           1264
     RT
RL
DR
DR
DR
DR
DR
                                                                                                                                                                                                                                                                                        ô
                                                                                                                                                                                                                                                                                                                        qq
                                                                                                                                                                                                                                                                                                                                                             Οy
                                                                                                                                                                                                                                                                                                                                                                                          Db
                                                                                                                                                                                                                                                                                                                                                                                                                               \Omega
                                                                                                                                                                                                                                                                                                                                                                                                                                                                g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   QQ
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          δy
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         Db
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            δ
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               QΩ
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  οy
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   Ωp
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      qq
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             Qγ
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            QQ
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               Ω
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  Op
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     φ
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     gg
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             δy
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             qα
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  Ωy
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   qq
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          ò
```

9

```
1427 FASGEASRPESVAESVKDEAGKAESRRESIAKTHKDESSLDKAKEQESRRESLAESIKPE 1486
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              1487 SGIDEKSALASKEASRPESVTDKSKEPSRRESIAESLKAESTKDEKSAPPSKEASRPGSV 1546
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              1641 ADSTKDEKSPLTSKDISRPESAVENVMDAPFKETSRPESAVGSMKDESMSKEPSRRESVK 1700
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 1701 DGAAQSRETSRPASVAESAKDGADDLKELSRPESTTQSKEAGSIKDEKSPLASE--EASR 1758
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  1759 PASVAESVKDEAEKSKEESRRESVAEKSPLPSKEASRPASVAESIKD-----EAEKS 1810
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              1811 KEESRRESV-----SRRESV 1856
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 1378 TEKSPLASKETSRPESATGSVKEDTEQTKSKKSPVPSR------PESEAKDKKSP 1426
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                1547 VESVKDETEKSKEPSRRESIAESAKPPIEFREVSRPESVIDGIKDESA-----KPESRRD 1601
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  1857 AEKSPLP----SKEASRPASVAESIKDEAEKSKEESRRESVAEKSPLPSKEASRPAS 1909
                                                                                                                                                                                                                    Drosophila melanogaster (Fruit fly).
Eukaryota; Metazoa; Arthropoda; Tracheata; Hexapoda; Insecta;
Pterygota; Neoptera: Endopterygota; Diptera; Brachycera; Muscomorpha;
Ephydroidea; Drosophilade; Drosophila.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  -----MAIEGLARMERRLQRATKRQKKQLEEDGIPVPSVVSRFPTAPYHH------ 132
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                133 -KSTNAEEREAKEP-----VLKTHSKDVE-REAE--IGVDDVVKMEPAATNIIEPEDAQD 183
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 184 A--AERGAARPPAVNSSYLPLPWKGRLGYACLNTYLRNAK-PPIFSSRTCRMASIV--- 236
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                237 -----DHRHPLQFED--EPEHHLKNKPD----KSKEPQDELGHKFVQEL------ 274
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    275 -GLANARDIVK--MLCWNEKYGIRFLR-------LSSEMFPFASHPVHGYK 315
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    376 PEQQNRDAVMIIHMGGQFGDKAATLERFKRNYARLSQSCKNRLVLENDDVGWTVHDLLPV 435
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   316 LAPFASEVLAEAGRVAAELGHRLTTHPGQFTQLGSPRKEVVESAIRDLEYHDELLSLLKL 375
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                436 CEELNIPMVLDYHHHNICFDPAHLREGTLDISDPKLQERIANTWKRKGIKQKMHYSEPCD 495
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                Ouery Match 4.5%: Score 155.5; DB 5; Length 5327; Best Local Similarity 19.4%; Pred. No. 0.28; Matches 151; Conservative 104; Mismatches 306; Indels 219; Gaps
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              496 GAVIPRD-----RRKHRPRVMTLPPCPP-----DMDLMIEAKDKEQAVFELMRTFKLPG 544
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   -----ELAGRAMGKD---ANGH----CLREGKEQE-----EGVK-- 87
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    3 SRKSKAAALDTPQSESSTFS-----STLDSSAPSPARNLRRSGRNILQPSSE-KDRDHE 55
                                                                                                                                                                                                                                                                                                                                                                          SEQUENCE FROM N.A. Papaga L., Siden-Kiamos I., Louis C.; Papagiannakis G., Spanos L., Siden-Kiamos I., Louis C.; "Sequencing the distal X chromosome of Drosophila melanogaster."; Submitted (JUL-1998) to the EMBL/GenBank/DDBJ databases.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  Submitted (AUG-1998) to the EMBL/GenBank/DDBJ databases. EMBL; AL031128; CAA20006.1; -. Flybase: Fspn0015390; futsel. SEQUENCE 5327 AA; 575942 MW; FEFEE23A118FF38A CRC64;
                                                                                                                      01-NOV-1998 (TrEMBLrel. 08, Last sequence update) U-UNN-2001 (TrEMBLrel. 17, Last annotation update) EG:49E4.1 PROPERIN PROPERIN OR CG:3064.
                                                    PRT; 5327 AA.
                                                                                                 01-NOV-1998 (TrEMBLrel. 08, Created)
                                                    PRELIMINARY;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   SEQUENCE FROM N.A.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   56 KRSGE-----
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                Benos P
                                                                      076891;
                                                 076891
RESULT 10
                       076891
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                q
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         g
                                                    ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  δ
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                Qγ
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              Pp
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                δy
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        QQ
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               δŏ
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              a
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                οy
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  δλ
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 ŏλ
```

```
40;
                                                                                                                                                                           2022 - AESIKDEAEKSKEESRRESV------AEKSPLPSKEASRPASVAESIKDEAE 2067
1910 VAESIKDEAEKSKEESRRESVAEKSPLPSKEASRPASVAESIKDEAEKSKEESRR---- 1964
                                                                                       1965 -ESVAEKSPLPSKEASRPASVAESIKDEAEKSKEESRRESVAEKSPLPSKEASRPASV-- 2021
                                                                                                                                   594 AADDVKDAFEGPKEVPEEERAMGGPYNRVYWPLGCEEWLKPKKREVKKGKVPEEVEDEGE 653
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            86; Mismatches 232; Indels 293; Gaps
                                            545 FEKINDMVPYDRDDENRPAPPV----KAPKKKKGGKRKRTTD-----EEAAEPEEVDT 593
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 379 ISIKQESADVSLVGNKELVTEEDL---DVFKQAQELSWEKIECESGVE----DCGRYPSV 431
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        491 YRRKDLSVFEVDGNMSKIYCQNLCLLAKLFLDHKTLYYDVEPFLFYVLTKNDEKGCHLVG 550
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         80 KEQEEGVKMAIEGLARMERRLQRATKRQKKQLEEDGIPVPS---VVSRFPTAPYHHKSTN 136
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 319 RPKKKGRKLLHEKAAQIKRRYAKPIGRPKNKLKQRLLSVTSDEGSMSAFTGRGSPDTDIK 378
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                137 AEEREAKEFVLKTHSKDVEREAEIGVDDVVKM-EPAATNIIEPEDAQDAAERGAARPPAV 195
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       432 IEFGKYEIQTWYSSPYPQEYARLP-KLYLCEFCLKYMKSKNILLRHSKKCGWFHPPANEI 490
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            ---IFSSRTCRMASI-VDHR-----HPLQF-----EDEPEHHLKN 254
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                255 KPDKSKEPQDELGHKFVQELGLANARDIVKMLCWNEKYGIRFLRLSSEMFPFASHPVHGY 314
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           551 YFSKEK-----HQRQGF 578
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       315 KLAPFASEVLAEAGRVAAELGHRLTTHPGQFTQLGSPRKEVVE-----SAIRDL 363
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             579 ------GRFLIDFSYLLSRREG---QAGSPEKPLSDLGRLSYLAYWKSVI--L 620
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       Thomas T., Voss A.K., Chowdhury K., Gruss P.; Joerkopf, a MYST family histone acctyltransferase, is required for normal cerebral cortex development."; Submitted (JAN-2000) to the BMBL/GenBank/DDBJ databases.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Buteleostomi;
Mammalia; Butheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  4.4%; Score 152.5; DB 11; Length 1763; 20.0%; Pred. No. 0.1;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  1763 AA; 196879 MW; 3A570C624B8094B0 CRC64;
                                                                                                                                                                                                                                                                                                                                    01-OCT-2000 (TrEMBLrel. 15, Created)
01-OCT-2000 (TrEMBLrel. 15, Last sequence update)
01-JUN-2001 (TrEMBLrel. 17, Last annotation update)
HISTONE ACETYLFRANSFERASE QUERKOPF.
                                                                                                                                                                                                                                                                                            PRT; 1763 AA.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               MGD; MGI:1858746; Morf.
InterPro; IPR001386; Linker_histone.
LiterPro; IPR002717; MOZ_SAS.
InterPro; IPR001965; PHD.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              AF222800; AAF26744.1; -.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 Pfam; PF01853; MOZ_SAS; 1. Pfam; PF00628; PHD; 2. SMART; SM00526; H15; 1. SMART; SM00249; PHD; 2.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              153; Conservative
                                                                                                                                                                                                                                                                                            PRELIMINARY;
                                                                                                                                                                                                                                                                                                                                                                                                                                              Mus musculus (Mouse).
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       Best Local Similarity
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               SEQUENCE FROM N.A.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                NCBI_TaxID=10090;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    TISSUE=BRAIN;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           Transferase
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  SEQUENCE
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    Query Match
                                                                                                                                                                                                                                                                                                                 09JKX5;
                                                                                                                                                                                                                                                                                          09JKX5
                                                                                                                                                                                                                                                  RESULT 11
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 EMBL;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              Matches
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            224
                                                                                                                                                                                                                                                                       Q9JKX5
                                                                                                                                   ŏ
                                                                                                                                                                                                                                                                                                                 qq
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                δλ
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        qq
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    Qγ
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               q
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         Οÿ
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    qq
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       δŏ
```

```
14424 FKLKPKKHPKPAYSVQDEEEEQ 14445
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            632 -- LKPKKREVKKGKVPEEVEDE 651
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  PRELIMINARY;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  SEQUENCE FROM N.A.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        CG14896 PROTEIN.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                NCBI_TaxID=7227;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             STRAIN-BERKELEY
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               CG14896
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         Q9VEU8;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     Q9VEU8
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       13
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        486
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   540
                                                                                            55
                                                                                                                                                                                         103
                                                                                                                                          13895
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     RESULT
Q9VEU8
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             ДQ
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         qq
                                                                                                                                                                                                                                                                                                                                 QQ
                                                                                                                                                                                                                                                                                                                                                                                                                              pp
                                                                                                                                                                                                                                                                                                                                                                                                                                                                             Οy
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           qq
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       qq
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          ŏ
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       Dp
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         òγ
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      qq
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        Qγ
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   qq
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     õ
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  οp
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  Q
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 ŏ
                                                                                            Óλ
                                                                                                                                                                                                                                                                                      δλ
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             33;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      Drosophila melanogaster (Fruit fly).
Eukaryota; Metazoa; Arthropoda; Tracheata; Hexapoda; Insecta;
Pterryota; Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;
Ephydroidea; Drosophilidae; Drosophila,
                                     621 EYLYRHHERHISIKAI----SRATGMCPH-----DIATTLQHLHMIDRRDGREVIIRRE 670
                                                                                                                             539 TFKLPGFEKINDMV----- 572
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  573 ----KGGKRKRTTDEEAAEPEEVDTAA-----DDVKDAPEGPKEVPEEERAMGGPYN 620
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           Gaps
364 EY----HDELLSLLKLPEQQNRDAVMIIHMGGQFGDKAATLER---FKRNYARLSQSCKN 416
                                                                                                                                                                                                                                      712 EAEKEAERLMEQASCWEKEEQEILSSRVSSRQSSAKVOSKNKYLHSPERRPVAGERGQLL 771
                                                                                                                                                                                            469 --PKLQERI---ANTWKR--KGIKQKMHYSEPCDGAVTPRDRRKHRPR--------
                                                                                                                                                                                                                                                                                         ---VMTLPP--CPPDMDLMIEAKDKEQAVFELMR
                                                                                                                                                                                                                                                                                                                       172 ELSKESSEEEEEEEEEEEEEEEEESIQTSPPRLTKPQ----SVSIKRKRPFVVKKKR
                                                                                                                                                                                                                                                                                                                                                                                                                    Query Match
4.1%; Score 144; DB 5; Length 16215;
Best Local Similarity 18.9%; Pred. No. 7.1;
Matches 140; Conservative 104; Mismatches 266; Indels 232;
                                                                                            417 RLVLENDDVGWTVHDLLPVCEELNIPMVLDYHHHNICFDPAHLREGTLDISD---
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     PROSITE; PS00179; AA_TRNA_LIGASE_II_1; UNKNOWN_1.
PROSITE; PS50002; SH3; 1.
SEQUENCE 16215 AA; 1841509 MW; 242C8765E00F7603 CRC64;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             621 RVYWPLGCE----EWLKPKKREVKKG------KVPEEVEDE 651
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           940 RQVWPKGAKRGLSKWKQSKER--KTGFKLNLYTPPETPMEPEDQ 981
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  Last annotation update)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  Created)
Last sequence update)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     PRT; 16215 AA
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       EMBL, AJ211740; CAB93524.1; -.
FlyBase; FBG10003432; sls.
InterPro; IPR002106; AALTRA_ligase_II.
InterPro; IPR003598; Ig_c2.
InterPro; IPR003598; Ig_c2.
InterPro; IPR00360; Ig_like.
InterPro; IPR003606; Ig_MHC.
InterPro; IPR001452; SH3.
Pfam; PF00041; ig; 48.
Pfam; PF00041; ig; 48.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 SMART; SM00408; IGC2; 15.
SMART; SM00410; IG_like; 34.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  01-OCT-2000 (TrEMBLrel. 15, 01-OCT-2000 (TrEMBLrel. 15, 01-JUN-2001 (TrEMBLrel. 17,
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               SLS OR D-TITIN OR CG1915.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     PRELIMINARY;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   SM00326; SH3; 1
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     SEQUENCE FROM N.A.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   SMART;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        RESULT 12
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       DAT THE PARTY OF T
                                                                                                                                            g
                                                                                                                                                                                            ò
                                                                                                                                                                                                                                      g
                                                                                                                                                                                                                                                                                         δ
                                                                                                                                                                                                                                                                                                                                      ద
                                                                                                                                                                                                                                                                                                                                                                                        ò
                                                                                                                                                                                                                                                                                                                                                                                                                                g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             δ
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       qq
                                                                                            ò
```

```
14264 LSIKRQEEEYEEGEDIEEFVVSQQRKPKP------LQITEEDEEAYTVKKLKR 14310
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  14371 SEAADEAKIKI----IQDFDDGEEPIIEEIRDDEDTIDEVEEP---EEYFVEELPPDEVD 14423
                                                                                                                                                                                                                   13950 --KKKKKK-----APVPEVVEEPEAEFLVKPKTPVQEVTEBAKITKSKKPV---KEEEAA 13998
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 13999 AELKVTITEEIPTEPEVQEIIEEIEEIEEIEERPAEYVIEVKESQPEAVEDKEVSLPKK--- 14055
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 ------KPKAPIV-------EEPEAEITLKPKVKSEEVQEEAK 14085
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                14086 IVKKKPKKIDEVAVADELTV-------KVEEEVVP-----EPIVEE 14119
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            14120 EVIE--------EFEIKKKPKEPEPEDIV-----DAAIVKLKKPEPVDAD 14156
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          14311 RKQVDIPEYADVENVTFRARSTKTKEDVDQEFNIALDSYAEEEISMSGKVKLKKPIKKTF 14370
                                                                                  13842 KKPKPEIKEEPEAEVTV-----STPKPVEEVEATSSIAVIPEQPTEEEAADLKITIIE 13894
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      ----QKMHYSEPCD-GAVTPRDRRKHRPRVMTLPPCPPDMDLMIEAKDKEQAVFELMRT 539
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          FK----PPVKAPKKKKGKRKTTD 582
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              267 ----GHKFVQELGLANARDIVKMLCWNEKYGIRFLRLSSEMFPFASHPVHGYKLAPFASE 322
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   323 VLAEAGRVAAELGHRLTTHPGQFTQLGSPRKEVVESAIRDLEYHDELLSLLKLPEQQNRD 382
                                                                                                                                                                    EKRSGEELAGRMMGKDANGHCLREGKEQEE-----GVKMAIEGLARMERRLQR 102
                                                                                                                                                                                                                                                                                                                                   ATKROKKOLEEDGIPVPSVVSR-----FPTAPYHHKSTNAEEREAKEPVLKTHSKDVE 155
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   209 GYACLNTYLRNAKPPIFSSRTCRMASIVDHRHPLQFEDEPEHHLKNKPD-KSKEPQDEL- 266
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  383 AVMIIHMGGQFGDKAATLERFKRNYARLSQSCKNRLV-LENDDVGWTVHDLLPVCEELNI 441
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       Drosophila melanogaster (Fruit fly).
Eukaryota; Metazoa; Arthropoda; Tracheata; Hexapoda; Insecta;
Pterygota; Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;
Ephydroidea; Drosophilidae; Drosophila.
    54
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          MEDLINE=2019G06, PubMed=10731132;
Adams M.D., Celniker S.E., Holt R.A., Evans C.A., Gocayne J.D.,
Amanatides P.G., Scherer S.E., Li P.W., Hoskins R.A., Galle R.F.,
4 RKSKAAALDTPQSESSTFSSTLDSSAPSPARNLRRSGRNIL---OPSSEKDRD-----H
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   156 REAEIGVDDVVKMEPAATNII------EPEDAQDAAERGAARPPAVNSSYLPLPWKGRL
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      PMV-----LDYHHHNICFDPAH-LREGTLDISDPKLQERIANTWKRKGIK------
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              583 EEAAEPEEVDTAADDVKDAPEGPKEVPEEERAMGGPYNRVYWPLGCEEW------
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     01-MAY-2000 (TrEMBLrel. 13, Created)
01-MAY-2000 (TrEMBLrel. 13, Last sequence update)
01-MAY-2000 (TrEMBLrel. 13, Last annotation update)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           828 AA
```

```
RA George R.A., Lewis S.E., Richards S., Ashburner M., Henderson S.N.,
RA Burton G.G., Wortman J.R., Yandell M.D., Zhang O., Chen L.X.,
RA Brandon R.C., Rogers Y.-H.C., Balazej R.G., Champen M., Pefelffer B.D.,
RA Brill J.F., Agbayani A., An H.-J., Andfersor P.E. Miklos G.L.G.,
RA Barli J.F., Agbayani A., An H.-J., Andfersor P.E. Baldwin D.,
RA Ballew R.M., Basu A., Bernen B.P. Brandari D., Bolshakov S.,
RA Benckon K.Y. Benco P.V., Bernan B.P., Brandari D., Bolshakov S.,
RA Borkova D., Botchan M.R., Bouck J., Brokstein D., Botshakov S.,
RA Borkova D., Botchan M.R., Bouck J., Brokstein D., Botshakov S.,
RA Borkova D., Botchan M.R., Buulter H., Cadieu E., Center A., Chandra I.,
RA Burtis K.C., Busan D.A., Laulter H., Cadieu E., Center A., Chandra I.,
RA Durbin K.J., Cawley S., Dahlke C., Davomport L.B., Davies P.,
RA Durbin K.J., Evangelista C.C., Ferraz C., Perriacs S., Fleischmann W.,
RA Durbin K.J., Brangelista C.C., Ferraz C., Gelbart W.M., Glasser K.,
A Goor E., Gorrell J.H., Gu Z., Guan P., Harris M.,
RA Hartis N.L., Harvy D., Heiman T.J., Hernandez J.R., Houck J.,
A Jalail M., Kalush F., Karpen G.H., Ke Z., Kennison J.A.,
A Jalail M., Kalush F., Karpen G.H., Ke Z., Kennison J.A.,
A Jalail M., Kalush F., Karpen G.H., Ke Z., Kennison J.A.,
A Jalail M., Malush F., Warft C., Mcraitz S., Kulp D., Lai Z.,
A Jalail M., Malush F., Marnor M., Ju J. Li Z., Liang Y., Lin X.,
A Merkluo G. M. Ishina N.V., Mobarry C., Morris J., Wosherfi A.,
A Merkluo G. Milshina N.V., Mobarry C., Morris J., Wosherfi A.,
A Blazzolo M., Pitton M., Sunders R.D.C., Scheeler F., Shen H.,
A Spier E., Spradling A.C., Stapleton M., Stropski M.P., Smith T.,
A Spier E., Spradling A.C., Stapleton M., Stropski M.P., Smith T.,
A Spier E., Spradling A.C., Stapleton M., Strong R., Wang Z. Y., Wassarman D.A., Weinscor G. M., Wunder S., Wu D., Yang S., Yao Q. A.,
A Leng X., Wassarman D.A., Weinscor G. Chocker E., Wang A., Wang S., Zhan M., Zhong Y., Zhan M., Zhong Y., Shu B.,
B Science 287:2185-2195(2000).
B Well., Shong S., Shu B., Shu B., Shu B., Shu
```

40; 77 PKRRARKASAKETNTTSDKKEKTVDILAPVKE-----VEPSLSIKKNSDHSESKG 126 127 POEPAPVATLNISKEEQDPDNVPNVSSENMAPKEPKKQERPKRRVRKASEDIVEVNAAID 186 -----ERRLQRA----TKRQKKQLE-----EDGIPVPSVVSRFPTA 128 187 AHNPETALASVQEDHLNQADQEMPKKRPRKSAELMLANVEEHAKEDHIJERPKRGKKPSL 246 129 PYHHKSTNAEER-----KTHSKDVE- 155 307 VSTSGEEHLAQIEEATEPVTESEPAQSELL-PEEGEENKTRRGKKPTVDTDEGPIK--- 362 207 RLGYACLNTYLRNAKP-PIFSSRTCRMASIVDHRHPLQFEDE------PEHHLKNKP 256 257 DKSKEPQDELGHKFVQELGLANARDIVKMLCWNEKYGIRFLRLSSEM---FPFASHPV-H 312 4.1%; Score 143.5; DB 5; Length 828; 19.9%; Pred. No. 0.16; Live 99; Mismatches 279; Indels 273; Gaps -----REAEIGVDDVVKMEPAATNIIEPEDAQDAAERGAARPPAVNSSYLPLPWKG 206 405 RASMEPE----HK--DELTAESSSEVVELP--TAKTTRGRKPSVDMEATIPEKKPPSRR 456 60 ------EELAGRAMGKDANGHCLREGKEQE---EGVKMAIEGLARM---- 96 363 -----KTPAEEPVPLPSHSRRGRKAT-----EDEVLPTADLAVPKTKLRGR- 404 2 PSRKSKAAALDTPQSESSTFSSTLDSSAPSPARNLRRSGRNILQP--SSEKDRDHEKRSG 59 Query Match
4.1%
Best Local Similarity 19.9%
Matches 162; Conservative 97 δ q δλ Op δλ QQ δy qq Q Óλ ò

A

6

Οy	313 G	GYKLAPFASEVLAEAGRVAAELGHRLTTHPGQFTQLGSPRKEVVESAIRD 362
qa	457 G	:: : : : : GHIDLQDLPTEIASPLVDTSGSPSKA 51
Qy	363 LI	HDELLSLLKLPEQQNRDAV
qq	517 SI	: : : : SD-AEELTPRRREGRNLPRKNYEEAPDDEKPHSGLRRARKPAATKSLASK- 565
ο _γ	417 RI	47
0 0	22	AFERDEY IFSEY INGEPANSEDIE
q _Q		GKA 65
٥y	513	LPPCPPDMDLMIEAKDKEQAVFELMRTFKLPGFEKINDWVPYDRDD 558
qq	660 AN	II SEEPPAKKP
QY	559	ENRPAPPYKAPKKKGGKRKRTTDEEAAEPEEVDTAADDVKD 600
qo	717 VF	VRGAGGKTKVETETEKOPPIKKPRGGARTKTPSEEAPQEEEQVKKSAARSRAKGTKA 773
Qy	601 A-1	A-PECPKEVPEEERAMGGPYNRVYW 624
RESULT Q9H945	JLT 14 945	
AC F	Q9H945 Q9H945;	PRELIMINARY; PRT; 914 AA.
IO	01-MAR-	01-MAR-2001 (TrEMBLrel. 16, Created)
DI	01-MAR-	
DE	TRIPLET	LUISUZZ FIS, CLONE NIZRP3000753, WEAKLY SIMILAR TO NEUROFILAMENT PROTEIN.
် ပ	Homo sa Eukaryo	
0 0 0	Mammali NCBI Ta	Mammalia; Butheria; Primates; Catarrhini; Hominidae; Homo.
R S	[1]	
A A	SEQUENCISOGAÍ	E FROM N.A. T., Ota T., Hayashi K., Suqiyama T., Otsuki T., Suzuki Y.,
RA 0	Nishika	. ,
RA RA	Saito K	II. Nimara M., Warahabe M., Hiraoka S., Ishil S., Kawal Y., (Yamamoto J., Wakamatsu A., Nakamura Y., Naqahari K.,
RA F-F	Masuho "NFDO h	Y., Kanehori K.;
R.	Submitt	ed (AUG-2000) to the EMBL/GenBank/DDBJ databases.
SQ	EMBL; A SEQUENC	EMBL; AK023084; BAB14396.1; SEQUENCE 914 Aa; 100583 MW; 7E7A9B3A5BAA53C2 CRC64;
Q B Z	ery Matc est Local	Ç
ÓΣ	11 LD	STLDSSAPSPARNLRRSGRNILQPSSEKDR-DHEKRSGEE-LA 63
qq	100 VD	VDPLEKESSTFTDENPSETEESEAAGGIGKLEGEDGDV-KCLSEKDTYDTSIDSLEENLD 158
QY	64 GR	GRMMGKDANGHCLREGKEQEEGVKMAIEGLARMERRLQRATKRQKKQLE 112
qq	: 159 KK	: : :
QY	113 ED	RFPTAPYHHKSTNAEEREAKEPVLKTHSKDVEREAEIGVI
qq	219 PG	PGGTTPKGLLPDSFHTETLEDTEKPSVNERPSEKDAEQEGEGGGIITRQTL 269
QY	173 TN	TNIIEPEDAQDAAERGAARPPAVNSSYLPLPWKGRLGYACLNTYLRNAKPPIFSSRTCRM 232

σ

```
------PKKKKGGKRKRTTDEEAAEPEEVDTAADDVKDAPE----GPKEVPEEERAM 615
                                                                                                                                                                                                    576 KYAPSSLYPSSLQAE--RLKKHLKKFPGATPAKNNWKMQKLWAKFRENPDQVEPEDGSDV 633
                                                                                                                                                                                                                                                                                                                                                                                                                                                                           ------ 518
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             634 SPGPNSEDSIEEVKEDRNSHPPANLPTPASTRILRKYSNIRGKLRAQQRLIKNEKMECPD 693
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 DMDLMIEAKDKEQAVFELMRTFKLPGFEKINDM----VPYDRDDENRPAPPVKA----- 568
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   694 --ALAVESKPSRKSVC-----INPLMSPKLALQVDADGFPVKPKSTEGMKGR 738
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            739 KGKQVSEILPKAEVQSKRKRT---EGSSPPDSKNKGPTVKASKEKHADGATKTPAAKRP- 794
                                                                                                                              336 HRLTTHPGQFTQLGSPRKE-----VVESAIRDLEYHDELLSLLKLPEQQNRDAVMIIHMG 390
                                                                                                                                                                                                                                                      391 GQF------GDKAATLERFKRNYARLSQSCKNRLVLENDDVGW 427
                                                                                                                                                                                                                                                                                        473 AQYMKVQKGWIQLEKEGQPTPRARNKSDKLKEIWKSKKR----SRKCRSSLESQK---- 523
                                                                                                                                                                                                                                                                                                                                                                                                  454 -FDPAHLREGTLDISDPKLQERI------ANTWKRKGIKQKMHYS----EPCDGA-V 498
270 KNMLDKEVKGLRGEIFPSRDP-ITTAGQPLPGERLEIYVQSKMDEKNAHIP-----S 320
                                  ASIVDHRHPLQFEDEPEH-------HLKNKPDKSKEPQDELGHKFVQELGL 276
                                                                                                         ANARDIVKMLCWNEKYGIRFLRLSSEMFPFASHPVHGYKLAPF-ASEVLAEAGRVAAELG 335
                                                                      ESIACKRDPEQAKEEPGHIPTQHVEEAVNEVDNENTQQKDDESDAPCSSLG-----LSS
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Hyperoartia;
Petromyzontiformes; Petromyzontidae; Petromyzon.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         SEQUENCE FROM N.A.
TISSUB-CENTRAL NERVOUS SYSTEM;
MEDLINE-95287814; PubMed=7770000;
Jacobs A.J., Kamholz J., Selzer M.E.;
Jarobs A.J., Romplor our Selzer M.E.;
mithe single lamprey neurofilament subunit (NF-180) lacks
multiphosphorylation repeats and is expressed selectively in
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        PRINTS; PR00194; TROPOMYOSIN.
PROSITE; PS00226; IF; UNKNOWN_1.
SEQUENCE 1110 AA; 123817 MW; 6558DA73DAF6974C CRC64;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        01-NOV-1996 (TrEMBLrel. 01, Created)
01-NOV-1996 (TrEMBLrel. 01, Last sequence update)
01-JUN-2001 (TrEMBLrel. 17, Last annotation update)
NF-180.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          PRT; 1110 AA
                                                                                                                                                                                                                                                                                                                                                                                                                                                                             ---RDRRKHRPRVMTLP-----
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   projection neurons.";
Brain Res. Mol. Brain Res. 29:43-52(1995)
EMBL; U19361; AAA80106.1; -.
InterPro; IPR000533; Tropomyosin.
InterPro; IPR001664; IF.
Pfam; PF00038; filament; 1.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               616 GGPYNRVYWPLGCEEWLKPKKREVKKGKV 644
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 795 -AARDRSSOP-----PKKTSLKENKV 814
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  Petromyzon marinus (Sea lamprey).
                                                                                                                                                                                                                                                                                                                               428 TVHDLLPVCEELNIPMVLDYHHHNIC---
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              PRELIMINARY;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       NCBI_TaxID=7757;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              091255
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          RESULT 15
                                                                                                                                                                                                                                                                                                                                                                   524
                                                                         321
                                                                                                           277
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              q
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               Q
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   q
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 g
                                                                                                                                                                                                                                                                                                                                                                 g
                                                                                                                                                                                                                                                                                                                                                                                                       ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                             δý
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       ð
                                                                                                                                                                                                                  q
                                                                                                                                                                                                                                                       δ
                                                                                                                                                                                                                                                                                    g
                                                                                                                                                                                                                                                                                                                             ò
                                       ò
                                                                    g
                                                                                                           δ
                                                                                                                                             Db
                                                                                                                                                                              φ
```

```
24;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     963 КРАRРКЕЕРЕDKAEPAKEKHSPVEERKPIKEIAKPAKAAPAKADKEPEA------- 1011
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             525 EAKDKEQAVFELMRTFKLPGFEKINDMVPYD------RDDENRPAPPVKAPKKK 573
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         574 GGKRKRTTDEEAAEP-----EEVDTAADDVKDAPEGPKEVPEEERAMGGPYNRVYWP 625
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           428 TVHDLLPVCEELNIPM--VLDYHHHNICFDPAHLREGTLDISDPKLQERIANTWKRKGIK 485
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       -----PTSEPEDEPKAEVVE-------KKGKAEAPKPKAKPAAAKK-EAKPVE 854
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                486 QKMHYSE-----PCDGAVTPRDRRKHRP---RVMTLPPCPPDMDLMI 524
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           915 EKEDDEEEE------EEVEEVKPEDAKPVKSKPAPAKEEEDEPKPAKQPPKPKR 962
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         262 --PQDELGHKFVQELGLANARDIVKMLCWNEKYGIRFLRLSSEMFPFASHPVHGYKLAPF 319
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       ASEVLAEAGRVAAELGHRLTTHPGQFTQLGSPRKEVVESAIRDI.EYHDELL.SLLKLPEQQ 379
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      AEEAEAEAEEEAA----KSEEEAAEEAKDEAEEEEAEEEAVEETEAATE------EAE 769
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 880 NRDAVMIIHMGGQFGDKAATLERFKRNYARLSQSCKNR------LVLENDDVGW 427
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 855 KEEEPEESPTEEEPKKPAAAKPAKAARAPKPRAAKAEAEEKPEPAKPAOAKPAPAAEEEED 914
                                                                                                                                                                                                                                                                                                                    549 EAAEBEEEAAEGEEEAEAEGEEAEEAEEVEEEAIEKAEAAEAKAEVEEEEAAEAEEEEEA 608
                                                                                                                                                                                                                                                                                                                                                                  143 KEPVLKTHSKDVEREAEIGVDDVVKMEPAATNIIEPEDAQDAAERGAARPPAVNSSYLPL 202
                                                                                                                                                                                                                                                                                                                                                                                                                                                               203 PWKGRLGYACLNTYLRNAKPPIFSSRTCRMASIVDHRHPLQFEDEPEHHLKNKPDKSKE- 261
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               -----KAKTQEAEVEEEBAEAAEAEBEBEBEBEBEBUBABSKEBEBEBSKEA 700
                                                                                                                                   429 PSAKPKAAKVEKKVVSKKPEIKVESEPISAQLDTDLEDLAQEEVMEAKAAPVVSAEKDEE 488
                                                                                                                                                                                 -KRSGEELAGRMMGKDANGHCLREGKEQEEGVKMAIEGLARMERRL 100
                                                                                                                                                                                                                             489 EEEEEEEKEEEEAEAEEEEEEDRGRKEGEAEAEEEAEEVEKEEAEEAEVEEAEAETEA 548
                                                                                                                                                                                                                                                                           101 QRATKRQK--------KQLEEDGI-PVPSVVSRFPTAPYHHKSTNAEEREA 142
                                                                                                                                                                                                                                                                                                                                                                                                  -----SD 721
                                                                                       2 PSRKSKAAALDTP-----QSESSTFSSTLDSSAPSPARNLRRSGRN1LQPSSEKDRD 53
DB 13; Length 1110;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            701 DAEEDEAEEEEVKE-----EEVTK----------
                        Best Local Similarity 17.6%; Pred. No. 0.31;
Matches 132; Conservative 102; Mismatches 275;
4.1%; Score 141.5; D
17.6%; Pred. No. 0.31;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           Search completed: January 15, 2002, 14:06:25 Job time: 964 sec
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   1012 -----AEPKKIEVKVKKVTKKVVEEIE 1033
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        626 LGCEEWLKPKKREVKKGKVPEEVEDEGE 653
                                                                                                                                                                                      54 HE---
    Query Match
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              815
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    652
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             722
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            qq
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             ŏ
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           qq
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         Οy
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     a
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   Qγ
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            q
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             ŏ
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  δò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 q
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            ζý
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        qq
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         ΩŸ
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    q_0
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   ŏ
                                                                                                                                                                                                                                                                                                                                                                                                                        QQ
                                                                                                                                                                                                                                                                                                                                                                                                                                                                        δλ
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  a
                                                                                                                                                                                                                                  QQ
                                                                                                                                                                                                                                                                                  οy
                                                                                                                                                                                                                                                                                                                           qd
                                                                                                                                                                                                                                                                                                                                                                            Qγ
```

•

H+-transporting AT 1	TED 9/KOO ALIGNMENTS RESULT 1 FOR STITUTE STATES	Predounclease - fission yeast (Schizosaccharomyces pombe) Paperles: Schizosaccharomyces pombe Paperles: Schizosaccharomyces Paperles: Schi	QY Z41 TLREGSLDLMPLIPTIRETWTRKGTTÖKQHYSESADPTAISGMKRRAHSDRVFDFPDPD 300
GenCore version 4.5 Copyright (c) 1993 - 2000 Compugen Ltd. OM protein - protein search, using sw model Run on: January 15, 2002, 13:52:22; Search time 74.94 Seconds (without alignments) 377.112 Million cell updates/sec Perfect score: 1979 Sequence: 1979 Sequence: 1 DDHAPREMFDCLDKPIPWRGKRLTARKRRSRKEEVEEDEK 371 Scoring table: BLOSUM62 Gapop 10.0, Gapext 0.5	Searched: 219241 seqs, 76174552 residues Total number of hits satisfying chosen parameters: 219241 Minimum DB seq length: 0	seq length: Ssing: Minimu Maximu Maximu Listin 1: Pir.68 1: Pir.68 1: Pir.68 1: Pir.68 3: Pir.68 4: Pir.69 4: Pir.69 5: Godery 6: Godery 7: Core Match L. 8 8 94: Pir.69 1:	97.5 4.9 487 2 T05679

á q ò pp

```
A;Cross-references: EMBL:249782; NID:9853752; PIDN:CAA89865.1; PID:9853757

R;Kunstr, F.; Ogasawara, N.; Moszer, I.; Albertini, A.M.; Alloni, G.; Azevedo, V.; Ber R;Kunstr, F.; Ogasawara, N.; Moszer, I.; Albertini, A.M.; Alloni, G.; Azevedo, V.; Brons, S.; Bron, S.; Broullet, S.; Brustra, C.V.; Caldwell, B.; Capuano, V.; Carter, N.M.; A.; Ehrlich, S.D.; Emmerson, P.T.; Entian, K.D.; Errington, J.; Pabret, C.; Ferrari, Nature 390, 249-256, 1997

A;Authors: Foulger, D.; Fritz, C.; Fujita, M.; Fujita, Y.; Fuma, S.; Galizzi, A.; Gal ecti, J.; Harwood, C.R.; Henaut, A.; Hilbert, H.; Holsappel, S.; Hosono, S.; Hullo, M Koetter, P.; Koningstein, G.; Krogh, S.; Kumano, M.; Kurita, K.; Lapidus, A.; Lardino Koetter, D.; Fordstein, G.; Krogh, S.; Kumano, M.; Kurita, H.; Masuda, S.; Mau A;Authors: Lauber, J.; Lazarevic, V.; Levine, A.; Liu, H.; Masuda, S.; Mau Keeter, M.; Rivolta, C.; Rocha, E.; Roche, B.; Rose, M.; Sadaie, Y.; Sanl Rieger, M.; Tamakoshi, A.; Tarpstra, P.; Tognoni, A.; Tosato, V.; Uchiya A;Authors: Schleich, S.; Schroeter, R.; Scoffone, F.; Sekiguchi, J.; Sakowska, A.; Tarpstra, P.; Tognoni, A.; Tosato, V.; Uchiya A;Authors: Yoshikawa, H.; Vamane, K.; Yasta, K.; Yasta, K.; Yoshida A;Title: The complete genome sequence of the Gram-Positive bacterium Bacillus subtili
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         R;Takami, H; Nakasone, K;; Takaki, Y;; Maeno, G;; Sasaki, R;; Masui, N; Fuji, F;; H Nucleic Acids Res. 28, 4317-4331, 2000
A;Title: Complete genome sequence of the alkaliphilic bacterium Bacillus halodurans a A;Reference number: A83650; MUID:20263314
R;Glaser, P.; Danchin, A. submitted to the EMBL Data Library, May 1995 A;Description: Cloning and sequencing of the Bacillus subtilis chromosomal region fro A;Reference number: $55414 A;Accasion: $55418
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            A;Cross-references: GB:299123; GB:AL009126; NID:92636240; PIDN:CAB15748.1; PID:e11862
A;Experimental source: strain 168
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        c;Species: Bacillus halodurans
C;Date: 01-Dec-2000 #sequence_revision 01-Dec-2000 #text_change 31-Dec-2000
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      hypothetical protein BH2597 [imported] - Bacillus halodurans (strain C-125)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           206 VPFVFDFHHFYANP----DDHADLNVALPRMIKTWERIGLOPKVHLSSPKSEQAI---- 256
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           90 QKEFREIGELVKTHQLKTSFHPNQFTLFTSPKESVTKNAVTDMAYHYRMLEAMGIADR-- 147
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    166 KDAVLIIHLGGTFEGKKETLDRFRKNYQRLSDSVKARLVLENDDVSWSVQDLLPLCQELN 225
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              226 IPLVLDWHHHNIVPGTLREGSLDLMPL1PTIRETWTRKGITQKQHYSESADPTAISGMKR 285
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           TIQRDGLESVKQLGTQNVLDLIKLVEWNHNFGIHFMRVSSDLFPFASH--AKYGYTLEFA 105
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         106 QSHLEEVGKLANKYNHRLTMHPGQYTQIASPREVVVDSAIRDLAYHDEILSRMKLNEQLN 165
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 --SVINIHIGGAYGNKDTATAQFHQNIKQLPQEIKERMTLENDDKTYTTEETLQVCEQED 205
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  Gaps
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         34 TERKEALLTVTKANLRNTM---RTLHYIIGHGIPLYRFSSSIVPLATHPDVMWDFVTPF- 89
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       A; Status: nucleic acid sequence not shown; translation not shown
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              Length 320;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           Indels
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   286 RAHSDRV---FDFPPCD-----PTMDLMIEAKEKEQAVFEL 318
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                257 RSHADYVDANFLLPLLERFROWGTNIDFMIEAKOKDKALLRL 298
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              61; Mismatches 109;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 18.2%; Score 360; DB 2; 29.8%; Pred. No. 1.5e-20;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  Conservative
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      Local Similarity
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            A; Molecule type: DNA
A; Residues: 1-320 <KUN>
                                                                                                                                                                                                                                                    A; Residues: 1-320 <GLA>
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          A; Status: preliminary
                                                                                                                                                                                       A;Status: preliminary
A;Molecule type: DNA
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         A; Molecule type: DNA
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   C;Accession: E83974
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        A; Accession: E83974
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               A; Accession: B70060
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     84;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               Query Match
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               A;Gene: ywjD
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          Best Loca
Matches
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         48
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         RESULT
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    QΩ
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             Óγ
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   qq
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   Qγ
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     ΟŽ
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               qq
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               Qγ
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         Db
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        C; Species: Neurospora crassa
C; Date: 10-Oct-1995 #sequence_revision 03-Nov-1995 #text_change 20-Jun-2000
C; Date: 10-Oct-1995 #sequence_revision 03-Nov-1995 #text_change 20-Jun-2000
R; Accession: S55262
R; Yajima, H: 7 Takao, M: Yasuhira, S.; Zhao, J.H.; Ishii, C.; Inoue, H.; Yasui, A.
EMBO J. 14, 2393-2399, 1995
A; Title: A eukaryotic gene encoding an endonuclease that specifically repairs DNA damage
A; Reference number: S55262; MUID:95292980
A; Accession: S5262
A; Status: nucleic acid sequence not shown
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   10;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      UV-endonuclease homolog ywjD - Bacillus subtilis
C;Species: Bacillus subtilis
C;Date: 15-Jul-1995 #sequence_revision 01-Sep-1995 #text_change 15-Oct-1999
C*Accession: S55418; B70060
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          321 SEVLAEAGRVAAELGHRLTTHPGQFTQLGSPRKEVVESAIRDLEYHDELLSLLKLPEQQN 380
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      166 KDAVLIIHLGGTFEGKKETLDRFRKNYQRLSDSVKARLVLENDDVSWSVQDLLPLCQELN 225
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           441 IPMVLDYHHHNICFDPAHLREGTLDISDPKLQERIANTWKRKGIKQKMHYSEPCD-GAVT 499
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  325
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      560 NRPAPPVKAPKKKKGGKRKRTTDEEAAEPEEVDTAADDVKDAPEGPKEVPEEERAMGGPY 619
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            106 QSHLEEVGKLANKYNHRLTMHPGQYTQIASPREVVVDSAIRDLAYHDEILSRMKLNEQLN 165
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    226 IPLVLDWHHHNIV--PGTLREGSLDLM--PLIPTIRETWTRKGITQKQHYSESADPTAIS 281
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              50 ---QRDGLESVKQLGTQNVLDLIKLVEWNHNFGIHFMRVSSDLFPFASHAKYGYTL-EFA 105
                                                                                               301 TMDLMIEAKEKEQAVFELCRRYELQNPPCPLEIMGPEYDQTRDGYYPPGAEKRLTARKRR 360
                                                                                                                                 469 TLREGSLDLMPLIPTIRETWTRKGITQKQHYSESADPTAISGMKRRAHSDRVFDFPPCDP 528
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      Gaps
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  98;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  48.5%; Score 960; DB 2; Length 656; 45.4%; Pred. No. 8.2e-67;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      Indels
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               A;Cross-references: EMBL:D11392; NID:g1526560; PID:g927215 C;Genetics: A;Gene: uvel
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                326 NPPCP----YDQTRDG----
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      54; Mismatches
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             345 ---YYPPGAEKRLTARKRRSRK----EEVEED 369
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       620 NRVYWPLGCEEWLKPKKREVKKGKVPEEVEDE 651
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   15 PIPWRGRLGYACLNTILRSMKERVFCSRTCRITTI--
                                                                                                                                                                                                                                                                                                                                                                                                RESULT 2
S55262
UV-endonuclease - Neurospora crassa
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           Best Local Similarity 45.4
Matches 205; Conservative
                                                                                                                                                                                                                                   361 SRKEEVEEDEK 371
                                                                                                                                                                                                                                                                         A; Residues: 1-656 <YAJ>
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   A; Molecule type: mRNA
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                Query Match
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 m
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 RESULT
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               S55418
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              pp
```

g

ολ

g δŽ g QY g ŏ Pp ŏλ g òλ

.. œ

```
Hi-transporting ATPase (EC 3.6.1.35) chain B, vacuolar - yeast (Saccharomyces cerevis NyAlternate names: Hi-transporting ATPase 57K chain; protein YBR1002; protein YBR127C C; Species: Saccharomyces cerevisiae C; Date: 26-Aug-1994 #sequence_revision 09-Sep-1994 #text_change 22-Jun-1999 C; Accession: S45996; S46570; S42165; A31487; A35681; B35681; A40178; S12281 F; Becam, A.M.; Herbert, C.J.; Masr, F.; Slonimski, P.P.; Zagulski, M. Submitted to the Protein Sequence Database, August 1994
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 A: Molecule type: DNA
A: Residues: 1-78, 'L',80-499,'TQEAPVRRTPAKKNL' < NEL>
A: Ross-references: BNBL:JO4450; NID:g171121; PIDN:AAA66890.1; PID:g171122
A: Cross-references: BNBL:JO4450; NID:g171121; PIDN:AAA66890.1; PID:g171122
T. Nelson, N.
J. Biol. Chem. 264, 1775-1778, 1989
A: Title: A conserved gene encoding the 57-KDa subunit of the yeast vacuolar H+-ATPase
A: Reference number: A31487; MUID:89109197
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   A; Experimental source: strain $288C
R; Becam, A.M.; Cullin, C.; Grzybowska, E.; Lacroute, F.; Nasr, F.; Ozier-Kalogeropoul
Yeast 10(Suppl.A), S1-S11, 1994
A; Title: The sequence of 29.7kb from the right arm of chromosome II reveals 13 comple
A; Reference number: $46569; MUID: 94378717
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 A.Cross-references: EMBL:235996; NID:9536412; PIDN:CAA85084.1; PID:9536413; MIPS:YBR1
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           Mol. Cell. Biol. 10, 3737-3749, 1990
A:Title: Role of vacuolar acidification in protein sorting and zymogen activation: a
A:Reference number: A35681; MUID:90287165
A:Accession: A35681
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               A; Status: nucleic acid sequence not shown; not compared with conceptual translation
                                                                                                               197 PDLPDPVRLRLGLENDERAVSPAELLPICEATGTPLVFDAHHHVV------HDKLPD 247
149 PEVRESSVRAMSAHARVMDGLGLARTPWN-----LLLLHGGKGGRGAE-----LAALI 196
                                                            193 QRLSDSVKARLVLENDDVSWSVQDLLPLCQELNIPLVLDWHHHNIVPGTLREGSLDLMP- 251
                                                                                                                                                                                         252 -LIPTIRETWTRKGIT-----QKQHYSESADPTAISGMKRRAHSDRVFDFPPCDPTMD- 303
                                                                                                                                                                                                                                                         248 QEDPSVREWVLRARATWQPPEWQVVHLS----NGIEGPQDRRHSHLIADFPSAYADVPW 302
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                A;Cross-references: EMBL:X75891; NID:g496856; PIDN:CAA53486.1; PID:g496858
A;Experimental source: strain S288C
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                R; Yamashiro, C.T.; Kane, P.M.; Wolczyk, D.F.; Preston, R.A.; Stevens, T.H.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                A;Residues: 1-78,'L',80-477,'ORFLMNFTIEPETMPTKMKKIRHKKLR' <NE2>
A;Cross-references: EMBL:J04450
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            R;Nelson, H.; Mandiyan, S.; Nelson, N.
submitted to the EMBL Data Library, January 1989
A;Reference number: S42165
A;Accession: S42165
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 A; Residues: 478-499, 'TQEAPVRRRTPAKKNL' <NEW>
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        R;Nelson, H.; Mandiyan, S.; Nelson, N.
J. Biol. Chem. 264, 5313, 1989
A;Reference number: A40178
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          A; Molecule type: DNA
A; Residues: 1-34;479-517 <YAM>
                                                                                                                                                                                                                                                                                                                         304 LMIEAKEKEQAVFEL 318
                                                                                                                                                                                                                                                                                                                                                                                  303 IEVEAKGKEEAIAAL 317
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                A; Cross-references: GB:J04450
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           A; Molecule type: protein
A; Residues: 'XV', 15-28 <YA2>
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        A; Molecule type: DNA
A; Residues: 1-517 <BEC>
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      A; Molecule type: DNA
A; Residues: 1-517 <BE2>
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     A; Molecule type: DNA
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             A; Molecule type: DNA
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          A;Accession: S45996
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 A; Accession: S46570
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          A; Accession: B35681
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             A; Contents: erratum
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        A; Accession: A40178
qq
                                                                                                                         q
                                                                                                                                                                                                                                                     qq
                                                                οy
                                                                                                                                                                                         ò
                                                                                                                                                                                                                                                                                                                     ŏ
                                                                                                                                                                                                                                                                                                                                                                                  q
             :Cross-references: GB:AP001516; GB:BA000004; NID:910175192; PIDN:BAB06316.1; GSPDB:GN00
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        C;Accession: C75350
R;White, O.; Eisen, J.A.; Heidelberg, J.F.; Hickey, E.K.; Peterson, J.D.; Dodson, R.J.; M.; Shen, M.; Vamathevan, J.J.; Lam, P.; McDonald, L.; Utterback, T.; Zalewski, C.; Ma Science 286, 1571-1577, 1999
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                A;Cross-references: GB:AE002022; GB:AE000513; NID:g6459590; PIDN:AAF11370.1; PID:g645959
A;Experimental source: strain Rl
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          A; Title: Genome sequence of the radioresistant bacterium Deinococcus radiodurans R1. A; Reference number: A75250; MUID:20036896
                                                                                                                                                                                                                                                                                    6
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        10;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  C;Species: Deinococcus radiodurans
C;Date: 03-Dec_1999 #sequence_revision 03-Dec-1999 #text_change 31-Mar-2000
                                                                                                                                                                                                                                                                                                                                                                                                                                                                      74 WNHNFGIHFMRVSSDLFPFASH---AKYGYTLEFAQSHLEEVGKLANKYNHRLTMHPGQY 130
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   58 HNLAYQISFFRLSSKLVPLVNHPLTEGWKYELAIAE-ELQAVGEFASEHQMRIDFHPDHF 116
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   131 TQIASPREVVVDSAIRDLAYHDEILSRMKLNEQLNKDAVLIIHLGGTFEGKKETLDRFRK 190
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            117 VVLNSEAKBITRRSLOTLLYHYKLLKGMEIDPRHR----CVLHVGGKKKGVEAGLEOFIE 172
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                191 NYQRLSDSVKARLVLENDDVSWSVQDLLPLCQELNIPLVLDWHHHNIVPGTLREGSLDLM 250
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            251 PLIPTIRETWTRKGITQKQHYSESADPTAISGMKRRAH-----SDRVFDF-----PPCDP 300
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     228 ETWORIVATWEDSPLPVKIHLS----SPLSGEDDPRHHDYINADRFIAFLHEIGADAVD 282
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          138 EVVVDSAIRDLAYHDEI----LSRMKLNEQLNKDAVLIIHLGGTFEGKKETLDRFRKNY 192
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          81 HFMRVSSDLFP---FASHAKYGYTLEFAQSHLEEVGKLANKYNHRLTMHPGQYTQIASPR 137
                                                                                                                                                                                                                                                                                Indels 37; Gaps
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     Indels 45; Gaps
                                                                                                                                                                                                                                                                                                                                      19 RGRLGYACLNTILRSMKERVFCSRTCRITTIQR-----DGLESVKQLGTQNVLDLIKLVE 73
                                                                                                                                                                                                                                                                                                                                                                                                   2 RIOFGYVAMSMELANASP----SKTMTATQFEKIEDHEAGLRKLERIAKTNLHNCLRLLK 57
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              21 RLGYACLNTILRSMKERVFCSRTCRITTIQRDGLESVKQLGTQNVLDLIKLVEWNHNFGI 80
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          31 QLGLVCLTVGPEVRFRTVTLSRYRALSPAEREA - - KLLDLYSSNIKTLRGAADYCAAHDI 88
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              173 NTASIPKSLLSMIMLENDDKSYTIDDVLYLGEKLAIPVVLDIHHHHDVL----HRSKSLQ
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                probable UV damage endonuclease - Deinococcus radiodurans (strain R1)
                                                                                                                                                                                                                    DB 2; Length 322;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        DB 2; Length 326;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 11.6%; Score 229.5; DB 2; 27.9%; Pred. No. 2.2e-10; Live 42; Mismatches 140;
                                                                                                                                                                                                                                                                             75; Mismatches 123;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      -
::
                                                                                                                                                                                                             15.5%; Score 306.5; DB 2 26.6%; Pred. No. 2.2e-16;
                                                  A; Experimental source: strain C-125 C; Genetics:
A; Gene: BH2597
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   <del>..</del>
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          : :|||||:|: |:||:| : | 283 HLHVMIEAKKKDLALFQLMK 302
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         301 TMDLMIEAKEKEQAVFELCR 320
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              Query Match
Best Local Similarity 27.98
Matches 88; Conservative
                                                                                                                                                                                                                                                                                Conservative
                                                                                                                                                                                                             Query Match
Best Local Similarity
Matches 85; Conserv
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      A; Molecule type: DNA
A; Residues: 1-326 <WHI>
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           A; Status: preliminary
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          A; Accession: C75350
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   A; Map position: 1
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 A; Gene: DR1819
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   C; Genetics
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                g
                                                                                                                                                                                                                                                                                                                                          ò
                                                                                                                                                                                                                                                                                                                                                                                                      g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                   ó
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            QQ
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     8
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         ŏ
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     용
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    g
```

C:Genetics	Qy 77 NFGIHFMRVSSDLFPFASHAKYGYT
A;Gene: SGD:VMA2; ATPsv; VAT2 A;Cross-references: MIPS:YBR127c; SGD:S0000331	
A:Map position: 2R A;Note: YBR127C C;Superfamily: H+-transporting ATP synthase alpha chain; H+-transporting ATP synthase al C;Keywords: ATP biosynthesis; hydrogen ion transport; hydrolase; membrane-associated com	Oy 122 RLTMHPGOYTQIASPREVVV Db 155 RLALTTAEYLAYQTERHVLTILTD
F;211-382/Domain: H+-transporting ATP synthase alpha chain homology <atp></atp>	
Query Match 5.6%; Score 111.5; DB 2; Length 517; Best Local Similarity 20.0%; Pred. No. 0.67; Matches 82; Conservative 62; Mismatches 140; Indels 125; Gaps 18;	Db 215 ERAĞRVEGRNGSITQIQy 233 HHHNIVPGTLREGSLDLMPLIPTIF
QY 27 LNTILRSMKERVFCSRTCRITIQRDGLESVKQLGTQNVLDLIKLVEWNH 76 : : : : : : : :	Db 263 HNKGIYPPINVLPSLSRLM Ov 280 ISGMKRRAHSDRVFDFPPCDPTMDI
77 NFGIHFMRVSSDLFPFASHAKYGYTLEFAQSHLEEVGKLANKYNH	317
: : :	Qy 321 RY ELQNPPCPLEIMGPEYDQTRI
QY 122 RLTMHPGQYTQIASPREVVVDSAIRDL-AYHDEILSRMKLNEQLNKDAVLII 172	365
QY 173 HLGGTFEGKKETLDRFRKNYQRLSDSVKARLVLENDDVSWSVQDLLPLCQELNIPLVLDW 232	RESULT 8 T44675 H+-transporting ATP synthase, chain N;Allernate names: V-ATPase chain I
QY 233 HHHNIVPGTLREGSLDLMPLIPTIRETWTRKGITQKQHYSESADPTA 279	C:Species: Desulturococcus sp. A;Variety: strain SY C:Date: 21-Jan-2000 #sequence_revi: C;Accession: T44675
QY 280 ISGMKRRAHSDRVFDFPPCDPTMDLMIEAKEKEQAVFELCR 320	R;Shibui, H.; Hamamoto, T.; Yohda, Biochem. Biophys. Res. Commun. 234 A;Tille: The stabilizing residues & A;Reference number: JC5532; MUID:9
Qy 321 RYELQNPPCPLEIMGPEYDOTRDCYYPPGAEKRLTARKRRKRKE 365	A; Accession: T44675 A; Status: preliminary; translated A; Molecule type: DNA A; Residues: 1-465 <shi></shi>
RESULT 7 B4224 UL-Transmorting Ambass (FC 3 6 1 35) whath B isoform 2 - houting (fragment)	A,Cross-references: EMBL/U948/7, P. A,Experimental source: strain SY C,Superfamily: H+-transporting ATP F,183-354/Domain: H+-transporting
HT-LIGHDOLLUIP ATFASE (E. 3.0.1.3.) C;Species: Bos primigenius taurus (cattle) C;Date: 04-Mar-1993 #sequence_revision 18-Nov-1994 #text_change 22-Jun-1999 C;Accession: B42554 D:Dinonlo K · K immameto C · Adachi I · Magner R · Fornac M	Ouery Match 5.6%; Best Local Similarity 19.1%; Matches 71; Conservative 55
A; Tatpory N.; A 3696-3706, 1992 J. Blol. Chem. 267, 3696-3706, 1992 A; Title: Differential expression of the "B" subunit of the vacuolar H(+)-ATPase in bovin A; Reference number: A42254; MUID:92156101	QY 27 LNT1LRSMKERVFCSRTCRLTT : : : 1 : 1 Db 141 MYTLYRGCKLPTFSGSGLPHNM
A. Arcesson: preliminary A. Status: preliminary A. Molecule type: mRNA	λo
A;Residues: 1-415 <puo> A;Cross-references: GB_M83130; NID:g162698; PIDN:AAA30389.1; PID:g162699</puo>	Db 185AAMGITYEEANFFKK
A; Experimental Source: Drain A; Note: sequence extracted from NCBI backbone (NCBIN:82547, NCBIP:82558) C; Superfamily: H+-transporting ATP synthase alpha chain; H+-transporting ATP synthase algorians and C; Keywords: ATP biosynthesis; hydrogen ion transport: hydrolase; membrane-associated com	OY 129 QYTQIASPREVVVDSA : : : Db 236 EYLAFDYDMQVLVILTDMTNYAEA
F;109-280/DOMAIN: H+-Udnsporting Air Synthase alpha Chain homorogy vair	QY 180 GKKETLDRFRKNYQRLSDSVKARL
Query Match 5.6%; Score 110.5; DB 2; Length 415; Best Local Similarity 20.0%; Pred. No. 0.58; Matches 82; Conservative 62; Mismatches 140; Indels 125; Gaps 18;	Db 296 GKKGSITQMPIL Qy 240 GTLREGSLDLMPLIPTIRETWTRK
**************************************	Db 344PIDVLPSLSRLMKDGIGK
Db 59 MNSIARGQKIPIFSASGLPHNEIAAQICRQAGLVRPTKDVHDGHEE 104	Qy 287 AHSDRVFDFPPCDFTMDLMIEAKE

```
M.; Kagawa, Y.
, 341-345, 1997
and the functional domains in the hyperthermophilic
7320421
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        synthase alpha chain; H+·transporting ATP synthase ATP synthase alpha chain homology <ATP>
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  n B [imported] - Desulfurococcus sp. (strain SY)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       sion 21-Jan-2000 #text_change 18-Aug-2000
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              Score 110; DB 2; Length 465;
Pred. No. 0.75;
); Mismatches 153; Indels 92; Gaps
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           EKEQAVFE-LCRRYELQNPPCPLEIMGPEYDQTRDGY 345
                                                                                                                                                                                                           | : |||: : : ||
|----PILTMPNDDITHPIPDLTGYITEGQIFVDRQL 262
                                                                                                                                                                                                                                                                                                                                                    SMIEAKEK-----LCR 320
                                                                                                                                                                                                                                                                                                                                                                                AQIARQAKVLGEE-----184
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         HLEEVGKLANK-----YNHRLTMHPG 128
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                SFEETGALERAVLFLNLADDPAIERIITPRMALTVA 235
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             IRDL-AYHDEILSRMKLNEQLNKDAVLIIHLGGTFE 179
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    LREISAAREEVPGRRGYPGYMYTDLATIYERGGRVR 295
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 VLENDDVSWSVQDLLPLCQELNIPLVLDWHHHNIVP 239
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          GITQKQH----YSES---ADPTAISGMKRR 286
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              RETWTRKGITQKQH-----YS---ESADPTA 279
                                                                                                                                                                                                                                                                                                             MKSAIGEGMTRKDHGDVSNQLYAKYAIGKDAAAMKA 316
                                          TARFEKRDFEENGSLERTSLFLNLANDPTIERIITP 154
                                                                                        ---DSAIRDI,-AYHDEILSRMKLNEQLNKDAVLII 172
                                                                                                                                                                               OSVKARLVLENDDVSWSVQDLLPLCQELNIPLVLDW 232
TLEFAQSHLEEVG-----KLANK----YNH 121
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    RDGLESVKQLGTQNVLDLIKLVEWNHNFGIHFMRVS 86
                                                                                                                                                                                                                                                                                                                                                                                                                                                                         | : :|: :||
DDADEDEEDPUTRSSGKKKDASQEE 412
                                                                                                                                                                                                                                                                                                                                                                                                                                                O--GYYPPGAEKRLTARKRRSRKEE 365
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 from GB/EMBL/DDBJ
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                IDN: AAB64417.1
```

A;Introns: 8/3; 73/3; 130/3; 193/1; 294/3; 348/2; 435/2 C;Superfamily: H+-transporting ATP synthase alpha chain; H+-transporting ATP synthase C;Keywords: hydrolase F;208-379/Domain: H+-transporting ATP synthase alpha chain homology <atp></atp>	Ouery Matches Matches Oy 5 Oy 49 Oy 108 Oy 108 Oy 108 Oy 145 Oy 264 Oy 264	SULT 11 387 GMTRKDHSDVSNQLYAAYAMGKDALAMRAVVGVEALSOEDLLY 42 SULT 11 9395
Db 398 SETDRKYLKFADRFEREFIAQRYDEDRSIFETLDLGWELLAELPESELKRVRKEY 452 QY 346YPPGAEKR 353 Db 453 ILKYHPKYRKR 463	Figure 9 Leansporting ATPase (EC 3.6.1.35) 57K chain - Arabidopsis thaliana Species: Arabidopsis thaliana (mouse-ear cress) Species: Arabidopsis thaliana (mouse-ear cress) Accession: A1886 Manolson, M.F.; Ouellette, B.F.F.; Filion, M.; Poole, R.J. Bloi. chem. 263, 17984, 1998 Manolson, M.F.; Ouellette, B.F.F.; Filion, M.; Poole, R.J. Title: CDNA sequence and homologies of the "57-kDa" nucleotide-binding subunit of the Reference number: A1886; MUID:89053962 Reference number: A1886, MUID:89053962 Molecule type: mRNA Residues: 1-492 ckANA Resi	9 27 LAWILLASKREKNYECSRTCRITIONLESWOLGTOWNLDLIKLVEWNH 76 158 MNSIARGGKIPLESAAGLPHNEIAAOICRCAGLVKRLEKTVDLLEDH 204 97 77NEGIFFRENGESAGLPHNEIAAOICRCAGLVKRLEKTVDLLEDH 204 97 77NEGIFFRENGESAGLPHNEIAAOICREL 1 1 1 1 1 1 1 1 1 1

```
A; Description: Analysis of the gene encoding the 57 KD B-subunit of the vacuolar H+-A A; Reference number: $25167
A; Accession: $25167
A; Molecule type: mRNA
A; Residues: 1-490 < DOW>
A; Cross references: EMBL: X67839; NID:98809; PIDN:CAA48034.1; PID:98810
A; Genetics:
A; Genetics: FlyBase: Wha55
A; Cross references: EMPL: X67839; NID:98809; PIDN:CAA48034.1; PID:98810
C; Superfamily: H+-transporting ATP synthase alpha chain; H+-transporting ATP synthase C; Keywords: ATP biosynthesis: hydrolase
F; 210-381/Domain: H+-transporting ATP synthase alpha chain homology <ATP>
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        A;cross-references: GB:S61797; NID:g237458; PIDN:AAB20098.1; PID:g237459 C;Superfamily: H+-transporting ATP synthase alpha chain; H+-transporting ATP synthase C;Keywords: hydrolase F;214-385/Domain: H+-transporting ATP synthase alpha chain homology <ATP>
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                A;Title: Molecular cloning and characterization of the B subunit of a vacuolar H(+)-A A;Reference number: $18395; MUID:92027804
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              H+-transporting ATPase (EC 3.6.1.35) chain B, vacuolar - tobacco budworm C;Species: Heliothis virescens (tobacco budworm) C;Date: 19-Mar_1997 #sequence_revision 19-Mar-1997 #text_change 21-Jan-2000
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               333 1LTMPNDDTTHPPIDDLTGYITEGQIYVDRQLHNRQIYP-----PVNVLPSLSRLMKSAI 386
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 46 ITTIQRDGLESVKQLGTQNVLDLIKLVEWNHNFGLHFMRVSSDLFPFASHAKYGYTLEFA 105
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 188 -----QAGL--VKLPG-KSVLD-----DHTDNFAIVFAAMGVNM------ETARFF 224
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          106 QSHLEEVGKLANK-----D 142
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               285 EALREVSAAREEVPGRRGFPGYMYTDLATIYERAGRVEGRNGSITQI------P 332
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              RLVLENDDVSWSVQDLLPLCQELNIPLVLDWHHHNIVPGTLREGSLDLMPLIPTIRETWT 261
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         132 DFLDIQGOPINPWSRIYPEEMIQTGISAIDVMNSIARGOKIPIFSAAGLPHNEIAAQICR 191
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                46 ITTIQRDGLESVKQLGTQNVLDLIKLVEWNHNFGIHFMRVSSDLFPFASHAKYGYTLEFA 105
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         225 KODFEENGSMENVCLFLNLANDPTIERIITPRLALTAAEFLAYOCEKHVLVILTDMSSYA 284
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    143 SAJRDL-AYHDE1LSRMKINEQLNKDAVLIIHLGGTFEGKKETLDRFRKNYQRLSDSVKA 201
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   86; Gaps
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     128 DFLDIOGOPINPWSRIYPEEMIQTGISAIDVMNSIARGOKIPIFSAAGLPHNEIAAQICR 187
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           ---RGRLGYACL,NTILRSMKERVF------CSRTCR 45
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     Gaps
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      7 EMPDCLDKPI-PW-----CSRTCRACLNTILRSMKERVF------CSRTCR 45
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              5.2%; Score 103; DB 2; Length 494; 19.7%; Pred. No. 2.9;
                                                                                                                                                                                                                                                                                                                                                                                                                                               5.3%; Score 104; DB 2; Length 490; 19.7%; Pred. No. 2.4;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     ;; Pred. No. 2.9; 48; Mismatches 114; Indels
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           48; Mismatches 114; Indels
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            ::
--
--
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  Biochem. Biophys. 291, 92-99, 1991
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         = = =
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 Ouery Match
Best Local Similarity 19.7%
Matches 61; Conservative
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     7 EMFDCLDKPI-PW-----
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    Best Local Similarity 19.79
Matches 61; Conservative
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     R;Gill, S.S.; Ross, L.S.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   A; Molecule type: mRNA
A; Residues: 1-494 <GIL>
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           262 RKGITQKQH 270
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      387 GEGMTRKDH 395
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                A;Status: preliminary
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    C; Accession: S18395
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           A; Accession: S18395
                                                                                                                                                                                                                                                                                                                                                                                                                                                         Query Match
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                οp
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               ŏλ
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        Ŏλ
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     qq
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          δλ
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         qq
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        δλ
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            Db
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              δ
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   δ
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    QΫ́
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               qq
                                                                                                                                                                                                                                                                                                                            RESULT 12
T44310
H+-transporting ATP synthase (EC 3.6.1.34) beta chain [imported] - Thermococcus sp. (str. C; Species: Thermococcus sp.
C; Species: Thermococcus sp.
A; Variety: strain KI
C; Date: 21-3n-2000 #sequence_revision 21-Jan-2000 #text_change 18-Aug-2000
C; Accession: T44310
C; Date: 21-Jan-2000 #sequence_revision 21-Jan-2000 #text_change 18-Aug-2000
C; Accession: T44310
R; Lida, T; Kanai, S; Inatomi, K; Kamagata, Y; Maruyama, T.
Biochim. Biophys: Acta 1329, 12-17, 1997
A; Title: Alpha- and beta-subunits of a v-type membrane ATPase in a hyperthermophilic sul A; Reference number: 222748; MUID:98036007
A; Reference number: 222748; MUID:98036007
A; Residues preliminary; translated from GB/EMBL/DDBJ
A; Ross-references: EMBL:D88772; NID:92605626; PIDN:BAA23343.1; PID:92605628
A; Cross-references: EMBL:D88772; NID:92605626; PIDN:BAA23343.1; PID:92605628
A; Cross-references: Extain KI
C; Superfamily: H+-transporting ATP synthase alpha chain; H+-transporting ATP synthase alpha chain homology <ATP>
F; 181-352/Domain: H+-transporting ATP synthase alpha chain homology <ATP>
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     H+-transporting ATPase (EC 3.6.1.35) chain B, vacuolar - fruit fly (Drosophila melanogad
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           13;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      C;Species: Drosophila melanogaster
C;Date: 22-Mov-1993 #sequence_revision 10-Nov-1995 #text_change 22-Jun-1999
C;Accession: S2167
R;Dow, J.A.T.; Kaiser, K.
Sybmitted to the EMBL Data Library, July 1992
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            396 SETDRKYLKFADRFEREFVAQRYDEDRSIFETLDLGWELL-----AELPESELKRVRKEY 450
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    240 GTLREGSLDLMPLIPTIRETWTRKGITQKQH------YSES---ADPTAISGMKRR 286
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           287 AHSDRVFDFPPCDPTMDLMIEAKEKEQAVFE-LCRRYELQNPPCPLEIMGPEYDQTRDGY 345
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             87 SDLFPFASHAKYGYTLE---FAQSHLEEVGKLANK-------YNHRLTMHPG 128
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    129 QYTQIASPREVVV-----DSAIRDL-AYHDEILSRMKLNEQLNKDAVLIIHLGGTFE 179
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            234 EYLAFDYDMQVLVILTDMTNYAEALREISAAREEVPGRRGYPGYMYTDLATIYERAGRVR 293
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              180 GKKETLDRFRKNYQRLSDSVKARLVLENDDVSWSVQDLLPLQQELNIPLVLDWHHHNIVP 239
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      294 GKKGSITQM------PILTMPDDDITHPIPDLTGYITEGQIVLSRELHRKGIYP 341
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            342 -----PIDVLPSLSRLMKDGIGKGRTREDHPQLSQQLYAAYAEGRSLRDLVAVVGEEAL 395
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  139 MNTLVRGQKLPIFSGSGLPHNMLAAQIARQAKVLGEE--------EQFAVVF---- 182
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            183 -----AAMGITYEEANFFKKSFEETGAIERAVLFLNLADDPAIERIITPRMALTVA 233
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           Gaps
                                                    310 EKEQAVFELCRRYELQNPPCPLEI-----MGPEYDQTR-----DGYYPPGAEKRLTAR 357
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   27 LNTILRSMKERVFCSRTCRITTIQRDGLESVKQLGTQNVLDLIKLVEWNHNFGIHFMRVS 86
                                                                                  92;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                DB 2; Length 463;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           Mismatches 154; Indels
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            r macch 5.3%; Score 104; DB Local Similarity 18.9%; Pred. No. 2.2; nes 70; Conservative ceres.
                                                                                                                                                                                  358 KRRSRKEEVEED 369
                                                                                                                                                                                                                                                459 KKEAKPEEATEE 470
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        346 --- YPPGAEKR 353
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    451 ILKYHPKYRKR 461
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   Query Match
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   RESULT 13
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              Matches
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    S25167
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                qq
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              δ
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            ŏ
                                                                                                                                                                                                                                          QQ
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    δ
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       õ
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  셤
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 QΥ
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  Ω
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        Q
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                g
                                                              ð
                                                                                                                      pp
                                                                                                                                                                                  οy
```

```
13;
                                                                                                                                                                  337 ILTMPNDDITHPIPDLTGYITEGQIYVDRQLHNRQIYP-----PVNVLPSLSRLMKSAI 390
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         ----QAGL--VKIPG-KSVLD-----DHEDNFAIVFAAMGVNM------ETARFF 228
                            -----YNHRLTMHPGQYTQIASPREVVV-----D 142
                                                      229 KODFEENGSMENVCLFLNLANDPTIERIITPRLALTAAEFLAYQCEKHVLVILTDMSSYA 288
                                                                                   143 SAIRDL-AYHDEILSRMKLNEQLNKDAVLIIHLGGTFEGKKETLDRFRKNYQRLSDSVKA 201
                                                                                                                                            RLVLENDDVSWSVQDLLPLCQELNIPLVLDWHHHNIVPGTLREGSLDLMPLIPTIRETWT 261
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                46 ITTIQRDGLESVKQLGTQNVLDLIKLVEWNHNFGIHFMRVSSDLFPFASHAKYGYTLEFA 105
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                QSHLEEVGKLANK-----D 142
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           229 KODFEENGSMENVCLFLNLANDPTIERIITPRLALTAAEFLAYOCEKHVLVILTDMSSYA 288
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          SAIRDL-AYHDEILSRMKLNEQLNKDAVLIIHLGGTFEGKKETLDRFRKNYQRLSDSVKA 201
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 RLVLENDDVSWSVQDLLPLCQELNIPLVLDWHHHNIVPGTLREGSLDLMPLIPTIRETWT 261
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    86; Gaps
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              7 EMFDCLDKPI-PW-----CSRTGYACLNTILRSMKERVF------CSRTCR 45
                                                                                                         EALREVSAAREEVPGRRGFPGYMYTDLATIYERAGRVEGRNGSITQI------P
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      5.2%; Score 103; DB 2; Length 494; 19.7%; Pred. No. 2.9; tive 48; Mismatches 114; Indels
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     289 EALREVSAAREEVPGRRGFPGYMYTDLATIYERAGRVEGRNGSITQI----
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              Search completed: January 15, 2002, 13:52:23
Job time: 392 sec
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   Query Match
Best Local Similarity 19.7%,
Matches 61; Conservative
                           106 QSHLEEVGKLANK----
                                                                                                                                                                                                  262 RKGITQKQH 270
                                                                                                                                                                                                                      :|:|:| |
391 GEGMTRKDH 399
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           262 RKGITQKQH 270
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           :1:1:1 | GEGMTRKDH 399
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                106
                                                                                                              583
                                                                                                                                         202
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    192
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          143
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    391
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  qq
                                                                                                                                       δ
                                                                                                                                                                      a
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               ŏ
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          qq
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                ŏ
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          q
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   q
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           q
                         ò
                                                    q
                                                                                 δ
                                                                                                         Db
                                                                                                                                                                                                  oy
od
```

,		·
	·	•
		•

```
January 15, 2002, 14:06:13; Search time 134.53 Seconds (without alignments) 403.382 Million cell updates/sec
GenCore version 4.5
Copyright (c) 1993 - 2000 Compugen Ltd.
                                                                                                      OM protein - protein search, using sw model
                                                                                                                                                       Run on:
```

1 DDHAPREMFDCLDKPIPWRG......KRLTARKRRSRKEEVEEDEK 371 Total number of hits satisfying chosen parameters: 473505 seqs, 146272329 residues BLOSUM62 Gapop 10.0 , Gapext 0.5 US-09-724-296-4 1979 Title: Perfect score: Scoring table: Sequence: Searched:

Post-processing: Minimum Match 0% Maximum Match 100% Listing first 45 summaries Minimum DB seq length: 0 Maximum DB seq length: 200000000

473505

sp_vertebrate:*
sp_unclassified:* sp_human:* sp_invertebrate:* sp_organelle:* sp_phage:* sp_archea:*
sp_bacteria:* sp_rodent:* sp_mammal:* sp_plant:* sp_virus:* SPTREMBL_17:* sp_fungi:* sp_mhc:* Database :

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

	Description	01000000000000000000000000000000000000	Q10908 SCH120SACCH	Oglafo megrospora	091409 bacilius ce						Ogufag carcing			Obusta parametrum	Obsept oncorny	Odigas organ	Obsobby Joseph S	Obnut Tentinua e	OO+200 COOKERS	P96018 staphylococ
SUMMARIES	ID	010988	001408	09L4F0	09L4E9	Q9K9P8	6M0S60	O9RTE6	O9XYC8	044427	09U6A8	096066	09GR11	09W6м4	09ASE0	0918A2	099982	09PIIK7	097792	P96018
	DB		m	~	7	7	7	7	Ŋ	2	2	S	'n	13	10	13	~	13	5	7
	Duery Match Length DB	599	929	317	325	322	305	326	496	492	489	509	881	502	488	496	588	503	1145	999
dР	Ouery Match	100.0	48.5	18.1	18.1	15.5	11.6	11.6	5.6	5.5	5.5	5.3	5.3	5.2	5.2	5.1	5.1	5.1	5.1	5.1
	Score	1979	096	359	359	306.5	229.5	229.5	111	109	108	105	104	103	102.5	101.5	101.5	101	101	100.5
	Result No.	П	7	ო	4	S	9	7	8	σ	10	11	12	13	14	15	16	17	18	19

Charle 77,000				052231 Jactococcus	054680 Jactococcus	O9sznl arabidonsis	O9lor5 arabidoosis	m/ syem edz Zugs60	09ln19 arabidonsis	P94264 borrelia bu	TILLISIUM SIIM 820800	O66172 agrobactori	OgraBB agrobactori	Obeska miscell	054272 stanbulococ	004650 arahidonsis	P73759 synechocyst	055455 synechocyst	099k16 mus musculu	ulusum sum 69m60	0916w3 pseudomonas	Ogsford chlamydia n	0929a2 chlamydia n	081507 arabidonsis	
0.09.52.77	3 09YHD1	0 091.ML4	1 099КАЗ	052231	054680	.0 09SZN1	.0 09LOR5		10 Q9LN19	P94264	1 099P78	066172		1 09ESK9	2 054272	_	Δ,		11 Q99K16		_	Q9S6P3	092962	0 081507	Q9PMJ1
635 1	511 1	507	400 3	383 2	383 2	487 1	520 1		485 1	554 2	634 1		822 2						633 1					512 1	519 2
5.0	5.0	5.0				4.9		4.9			4.9		4.9	4.9	4.9	4.9	4.9	4.9	4.8	4.8	4.8	4.8	4.8	4.7	4.7
99.5	66	98.5	98	97.5	97.5	97.5	97.5	97.5	26	97	97	67	97	97	96.5	96.5	96	96	95	92	95	94.5	94.5	94	94
20	21	22	23	24	25	56	27	28	59	30	31	32	33	34	35	36	37	38	39	40	4.1	42	43	44	45

ALIGNMENTS

```
Takao M., Yonemasu R., Yamamoto K., Yasui A.; "Characterization of a UV endonuclease gene from the fission yeast Schizosaccharomyces pombe and its_bacterial homolog."; Nucleic Acids Res. 24:1267-1271 (1996).
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   Lyne M., Rajandream M.A., Barrell B.G., Lucas M., Gaillardin C.; Submitted (JUN-1998) to the EMBL/GenBank/DDBJ databases. EMBL: D78571: BAA11415.1; --
EMBL: U78487; AAC49664.1; --
                                                                                                                                                                                                                                                                                                                                                                          Davey S., Nass M.L., Ferrer J.V., Sidik K., Eisenberger A., Mitchell D.L., Freyer G.A.;
                                                                                                                                                                                                                                                                                                                                                                                                         "The fission yeast UVDR DNA repair pathway is inducible."; Nucleic Acids Res. 25:1002-1008(1997).
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              68816 MW; C09ED7219D5895BE CRC64;
                                                                                                                                  Schizosaccharomyces pombe (Fission yeast).
Eukaryota; Fungi; Ascomycota; Schizosaccharomycetes;
Schizosaccharomycetales; Schizosaccharomycetaceae;
                                                                           Last sequence update)
Last annotation update)
                                PRT; 599 AA.
                                                      01-NOV-1996 (TrEMBLrel. 01, Created)
01-NOV-1996 (TrEMBLrel. 01, Last seq
01-NOV-1998 (TrEMBLrel. 08, Last anno
                                                                                                                                                                                                                                                                                                                              STRAIN=SP972;
MEDLINE=96188860; PubMed=8614629;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               EMBL; AL023859; CAA19577.1;
                             PRELIMINARY;
                                                                                                                                                                               Schizosaccharomyces.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           Endonuclease.
SFOUENCE 599 AA;
                                                                                                                                                                                                                           SEQUENCE FROM N.A.
                                                                                                                                                                                                                                                                                                                                                                                                                                                       SEQUENCE FROM N.A.
                                                                                                    UV - ENDONUCLEASE.
                                                                                                                                                                                               NCBI_TaxID=4896;
                                          P87339;
                                                                                                                       UVDE OR UVE1+.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                       STRAIN=972H-;
                                           010988;
                             010988
RESULT
               010988
```

Query Match

100.0%; Score 1979; DB 3; Length 599;

```
10;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       "A eukaryotic gene encoding an endonuclease that specifically repairs DNA damaged by ultraviolet light."; EMBO 0. 14:393-2399(1995). EMBL: D1392; CAB20113.1: -.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             50 ---QRDGLESVKQLGTQNVLDLIKLVEWNHNFGIHFMRVSSDLFPFASHAKYGYTL-EFA 105
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      106 OSHLEEVGKLANKYNHRLTMHPGQYTQIASPREVVVDSAIRDLAYHDEILSRMKLNEQLN 165
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            301 IMDIMIEAKEKEQAVFELCRRYELQNPPCPLEIMGPEYDQTRDGYYPPGAEKRLTARKRR 360
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               181 KKETLDRFRKNYQRLSDSVKARLVLENDDVSWSVQDLLPLCQELNIPLVLDWHHHNIVPG 240
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               409 KKETLDRFRKNYQRLSDSVKARLVLENDDVSWSVQDLLPLCQELNIPLVLDWHHHNIVPG 468
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             241 TLREGSLDLMPLIPTIRETWTRKGITQKQHYSESADPTAISGMKRRAHSDRVFDFPPCDP 300
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   469 TLREGSLDLMPLIPTIRETWTRKGITOKQHYSESADPTAISGMKRRAHSDRVFDFPPCDP 528
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         Gaps
                                                                                                                                      61 GTQNVLDLIKLVEWNHNFGIHFMRVSSDLFPFASHAKYGYTLEFAQSHLEEVGKLANKYN 120
                                                                                                                                                                                                                                                                                                                                                                121 HRLTMHPGQYTQIASPREVVVDSAIRDLAYHDEILSRMKLNEQLNKDAVLIIHLGGTFEG 180
                                                                                                1 DDHAPREMFDCLDKPIPWRGRLGYACLNTILRSMKERVFCSRTCRITTIQRDGLESVKQL 60
                                  Gaps
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   rajima H., Takao M., Yasuhira S., Zhao J.H., Ishii C., Inoue H.,
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     .86
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   Eukaryota; Fungi; Ascomycota; Pezizomycotina; Sordariomycetes; Sordariales; Sordariaceae; Neurospora.
                                  .;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    48.5%; Score 960; DB 3; Length 656; 45.4%; Pred. No. 1e-63;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         95; Indels
                                  Indels
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             656 AA; 74413 MW; 4B231E7A8735C86F CRC64;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    01-NOV-1996 (TrEMBLrel. 01, Created)
01-NAY-1999 (TrEMBLrel. 10, Last sequence update)
01-NAY-1999 (TrEMBLrel. 10, Last annotation update)
   Pred. No. 5.1e-140;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       15 PIPWRGRLGYACLNTILRSMKERVFCSRTCRITTI------
                                      .;
0
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            54; Mismatches
                                      0; Mismatches
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          656 AA.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   SEQUENCE FROM N.A.
STRAIN=74-OR23-1A;
MEDLINE=95292980; PubMed=7774597;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          PRT;
100.08;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  EMBL; D11392; CAB20113.1;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            Matches 205; Conservative
                                            Conservative
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             PRELIMINARY;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                361 SRKEEVEEDEK 371
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              589 SRKEEVEEDEK 599
Best Local Similarity
Matches 371; Conserva
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             Best Local Similarity
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              Neurospora crassa.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    NCBI_TaxID=5141;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            UV - ENDONUCLEASE.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        Endonuclease
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              Query Match
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    SEQUENCE
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            001408;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             001408
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               RESULT
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              001400

10 017

10 017

10 017

10 017

10 017

10 017

10 017

10 017

10 017

10 017

10 017

10 017

10 017

10 017

10 017

10 017

10 017

10 017

10 017

10 017

10 017

10 017

10 017

10 017

10 017

10 017

10 017

10 017

10 017

10 017

10 017

10 017

10 017

10 017

10 017

10 017

10 017

10 017

10 017

10 017

10 017

10 017

10 017

10 017

10 017

10 017

10 017

10 017

10 017

10 017

10 017

10 017

10 017

10 017

10 017

10 017

10 017

10 017

10 017

10 017

10 017

10 017

10 017

10 017

10 017

10 017

10 017

10 017

10 017

10 017

10 017

10 017

10 017

10 017

10 017

10 017

10 017

10 017

10 017

10 017

10 017

10 017

10 017

10 017

10 017

10 017

10 017

10 017

10 017

10 017

10 017

10 017

10 017

10 017

10 017

10 017

10 017

10 017

10 017

10 017

10 017

10 017

10 017

10 017

10 017

10 017

10 017

10 017

10 017

10 017

10 017

10 017

10 017

10 017

10 017

10 017

10 017

10 017

10 017

10 017

10 017

10 017

10 017

10 017

10 017

10 017

10 017

10 017

10 017

10 017

10 017

10 017

10 017

10 017

10 017

10 017

10 017

10 017

10 017

10 017

10 017

10 017

10 017

10 017

10 017

10 017

10 017

10 017

10 017

10 017

10 017

10 017

10 017

10 017

10 017

10 017

10 017

10 017

10 017

10 017

10 017

10 017

10 017

10 017

10 017

10 017

10 017

10 017

10 017

10 017

10 017

10 017

10 017

10 017

10 017

10 017

10 017

10 017

10 017

10 017

10 017

10 017

10 017

10 017

10 017

10 017

10 017

10 017

10 017

10 017

10 017

10 017

10 017

10 017

10 017

10 017

10 017

10 017

10 017

10 017

10 017

10 017

10 017

10 017

10 017

10 017

10 017

10 017

10 017

10 017

10 017

10 017

10 017

10 017

10 017

10 017

10 017

10 017

10 017

10 017

10 017

10 017

10 017

10 017

10 017

10 017

10 017

10 017

10 017

10 017

10 017

10 017

10 017

10 017

10 017

10 017

10 017

10 017

10 017

10 017

10 017

10 017

10 017

10 017

10 017

10 017

10 017

10 017

10 017

10 017

10 017

10 017

10 017

10 017

10 017
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            Вb
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          ò
                                                                                                                                                                                                                                                                                                                                                                                                                                  Dp
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     q
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       δ
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   οy
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              qq
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   δλ
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              qq
                                                                                                                                                                         g
                                                                                                                                                                                                                                           ŏ
                                                                                                                                                                                                                                                                                                                                                                       δλ
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       ò
```

```
STRAINS—ATCC 14579 TYPE STRAIN;
STRAINS—ATCC 14579 TYPE STRAIN;
MEDLINE—20055637; PubMed=10589720;
Okstad O., Gominet M., Purnelle B., Rose M., Lereclus D., Kolsto A.B.;
Sequence analysis of three Bacillus cereus loci under PICR-regulated genes encoding degradative enzymes and enterotoxin.";
Microbiology 145:3129-3138 (1999).
EMBL; AJ243712; CAB69813.1; -.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    121 HRLTMHPGQYTQIASPREVVVDSAIRDLAYHDEILSRMKLNEQLNKDAVLIIHLGGTFEG 180
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           181 KKETLDRFFRNYQRLSDSVKARLVLENDDVSWSVQDLLPLCQELNIPLVLDWHHH--NIV 238
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              239 PGTLREGSLDLMPLIPTIRETWTRKGITQKQHYSESADPTAISGMKRRAHSDRVFDFPPC 298
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              63 QNVLDLIKLVEWNHNFGIHFMRVSSDLFPFASH--AKYGYTLEFAQSHLEEVGKLANKYN 120
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    46 ONLEHTIRILHYNIAHEIPLYRLSSSIVPLATHPEVEFDYIGVFTPL-WRRIGALIKEHN 104
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     221 EEPLEE-----LLPAIFETWSHTNISPKVHISSPR----SEKEFRAHAEYI-DLEFI 267
                                                                                                                                                                                  282 GMKRRAHSDRVFDFPPCDPTMDLMIEAKEKEQAVFELCRRYEL-------Q 325
                                                                                                                                                                                                        560 NRPAPPVKAPKKKKGGKRKRTTDEEAAEPEEVDTAADDVKDAPEGPKEVPEEERAMGGPY 619
                                                                    381 RDAVMIIHMGGQEGDKAATLERFKRNYARLSQSCKNRLVLENDDVGWTVHDLLPVCEELN 440
                                                                                                                                321 SEVLABAGKVAAELGHRLTTHPGQFTQLGSPRKEVVESAIRDLEYHDELLSLLKLPEQQN 380
                                                                                                            226 IPLVLDWHHHNIV--PGTLREGSLDLM--PLIPTIRETWTRKGITQKQHYSESADPTAIS 281
                                      166 KDAVLITHLGGTFEGKKETLDRFRKNYQRLSDSVKARLVLENDDVSWSVQDLLPLCQELN
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      Length 317;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             95; Indels
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    317 AA; 36990 MW; A6AC35F5800A22C3 CRC64;
                                                                                                                                                                                                                                                                -----LEIMGPE----YDQTRDG----
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     01-OCT-2000 (TrEMBLrel. 15, Last sequence update) 01-MAR-2001 (TrEMBLrel. 16, Last annotation update)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          Bacteria; Firmicutes; Bacillus/Clostridium group; Bacillus/Staphylococcus group; Bacillus.
NCBL_TaxID=1396;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        18.1%; Score 359; DB 2; 30.7%; Pred. No. 3.5e-19;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              Best Local Similarity 30.7%; Pred. No. 3.5e
Matches 83; Conservative 60; Mismatches
                                                                                                                                                                                                                                                                                                                                      345 ---YYPPGAEKRLTARKRRSRK----EEVEED 369
                                                                                                                                                                                                                                                                                                                                                                         620 NRVYWPLGCEEWLKPKKREVKKGKVPEEVEDE 651
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            299 DP-----TMDLMIEAKEKEQAVFEL 318
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               01-OCT-2000 (TrEMBLrel. 15, Created)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                    PRT;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              PUTATIVE UV-ENDONUCLEASE.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                    PRELIMINARY;
                                                                                                                                                                                                                                                                326 NPPCP-----
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   Bacillus cereus
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        Endonuclease
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              Query Match
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            SEQUENCE
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        09L4F0;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                    Q9L4F0
                                                                                                                                                                                                                                                                                                                                                                                                                                   RESULT
Q9L4F0
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           q
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 qq
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         a
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            δà
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             a
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     pp
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          δÀ
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                Öζ
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        ÖΣ
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    δà
                                                                                                                                                                                                                                                                                                   qq
                                                                                                                                                                                                                           qq
                                                                                                                                                                                                                                                                                                                                            ò
                                          Ωy
                                                                             q
                                                                                                                  õ
                                                                                                                                                     q
                                                                                                                                                                                          \stackrel{\circ}{\circ}
                                                                                                                                                                                                                                                                  Q
```

.. 8

```
42;
                                                                                                                                                                                                                                                                                                                                                                                           301 IMDLMIEAKEKEQAVFELCR 320
                                                                                                         85; Conservative
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    Conservative
                                                                                                                                                                                                                                                                                                                                                                                                                                                                   PRELIMINARY;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      Deinococcus radiodurans
                                                                                            Sest Local Similarity
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          Ouery Match
Best Local Similarity
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        SEQUENCE FROM N.A.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               NCBI_Tax1D=1299;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              UV - ENDONUCLEASE.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           Endonuclease.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 88;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  STRAIN=KR1;
                                                                                  Query Match
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       SEQUENCE
                                                                                                                                                                                                                                                                                                                                                                                                                                                                  6M0S60
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                09S0M9;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          UVSCDE
                                                                                                        Matches
                                                                                                                                                                                                                                                                                                                                                                                                                                            9
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  Matches
                                                                                                                                                                                                                                                                                                                                                                                                                                            RESULT
RT
DR
SO
SO
                                                                                                                                 δλ
                                                                                                                                                       O
                                                                                                                                                                                δŏ
                                                                                                                                                                                                      q
                                                                                                                                                                                                                              Qγ
                                                                                                                                                                                                                                                       Вb
                                                                                                                                                                                                                                                                               οğ
                                                                                                                                                                                                                                                                                                     qq
                                                                                                                                                                                                                                                                                                                               ŏ
                                                                                                                                                                                                                                                                                                                                                      qq
                                                                                                                                                                                                                                                                                                                                                                              ŏ
                                                                                                                                                                                                                                                                                                                                                                                                     qq
                                                                                                                                                                                                                                                                                                                                                                                                                                                                     Ωp
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                QQ
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          δλ
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          òλ
                                                                                                                                                                                                                                                                                                                                                     8
                                                                                                                                                                       SEQUENCE FROM N.A.
STRAINSATC 14579 TYPE STRAIN;
MEDLINE-20055637: PubMed=10589720;
OKSTAO O., Gominet M., Purnelle B., Rose M., Lereclus D., Kolsto A.B.;
"Sequence analysis of three Bacillus cereus loci under PICR-regulated genes encoding degradative enzymes and enterotoxin.";
Microbiology 145:319-3138(1999).
EMBL: AJ243712; CAB69814.1; -.
                                                                                                                                                                                                                                                                                                                                                                         63 QNVLDLIKLVEWNHNFGIHFMRVSSDLFPFASH--AKYGYTLEFAQSHLEEVGKLANKYN 120
                                                                                                                                                                                                                                                                                                                                                                                                  54 QNLEHTIRILHYNIAHEIPLYRLSSSIVPLATHPEVEFDYIGVFTPL-WRKIGALIKEHN 112
                                                                                                                                                                                                                                                                                                                                                                                                                          121 HRLTMHPGQYTQIASPREVVVDSAIRDLAYHDEILSRMKLNEQLNKDAVLIIHLGGTFEG 180
                                                                                                                                                                                                                                                                                                                                                                                                                                       181 KKETLDRFRKNYQRLSDSVKARLVLENDDVSWSVQDLLPLCQELNIPLVLDWHHH--NIV 238
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       229 EEPLEE-----LPAIFETWSHTNISPKVHISSPR-----SEKEFRAHAEYI-DLEFI 275
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        239 PGTLREGSLDLMPLIPTIRETWTRKGITQKQHYSESADPTAISGMKRRAHSDRVFDFPPC 298
                                                                                                                                                                                                                                                                                                                                                      Gaps
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               Takami H., Nakasone K., Takaki Y., Maeno G., Sasaki R., Masui N. Fuji F., Hirama C., Nakamura Y., Ogasawara N., Kuhara S., Horikoshi K.;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      "Complete genome sequence of the alkaliphilic bacterium Bacillus
                                                                                                                                                                                                                                                                                                                                                     32;
                                                                                                                                                                                                                                                                                                                            Length 325;
                                                                                                                                                                                                                                                                                                                                                     Indels
                                                                                                                                                                                                                                                                                     325 AA; 37900 MW; 07EFBDD1FA580653 CRC64;
                                                                              Last annotation update)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  Last sequence update)
Last annotation update)
                                                                                                                                                                                                                                                                                                                   Query Match
Best Local Similarity 30.7%; Pred. No. 3.6e-19;
Matches 83; Conservative 60; Mismatches 95;
                                                                                                                            Bacteria; Firmicutes; Bacillus/Clostridium group; Bacillus/Staphylococcus group; Bacillus.
                                                                  Last sequence update)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             Bacteria; Firmicutes; Bacillus/Clostridium group; Bacillus/Staphylococcus group; Bacillus.
                                325 AA
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                322 AA
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         299 DP-----TMDLMIEAKEKEQAVFEL 318
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       276 KPFLHIAKKHNHNFDIMIESKOKDLALFOL 305
                                                       Created)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 01-07T-2000 (TrEMBLrel. 15, Created)
01-0CT-2000 (TrEMBLrel. 15, Last seq
01-0CT-2000 (TrEMBLrel. 15, Last ann
                                PRT;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          SEQUENCE FROM N.A.
STRAIN-C-125 / JCM 9153;
MEDLINE-20512582; Pubmed-11058132;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               PRT;
                                                  01-OCT 2000 (TrEMBLEE) 15, 01-OCT-2000 (TREMBLEE) 15, 01-MAR-2001 (TREMBLEE) 16, PUTATIVE UV-ENDONUCLEASE.
                             PRELIMINARY;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              PRELIMINARY;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  Bacillus halodurans.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      NCBI_TaxID=86665;
                                                                                                                                                    NCBI_TaxID=1396;
                                                                                                                 Bacillus cereus
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        BH2597 PROTEIN.
                                                                                                                                                                                                                                                                          Endonuclease
                                                                                                                                                                                                                                                                                       SEOUENCE
                         Q9L4E9;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             Q9K9P8
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         09K9P8
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     S
     RESULT
                09L4E9
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     RESULT
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 09K9P8
                                          ó
                                                                                                                                                                                                                                                                                                                                                                                                g
                                                                                                                                                                                                                                                                                                                                                                                                                           ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                 g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                          ð
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        ö
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               a
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               qq
```

```
10;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  74 WNHNFGIHFMRVSSDLFPFASH---AKYGYTLEFAQSHLEEVGKLANKYNHRLTMHPGQY 130
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        131 TQIASPREVVVDSAIRDLAYHDEILSRMKLNEQLNKDAVLIIHLGGTFEGKKETLDRFRK 190
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    191 NYQRLSDSVKAREVLENDDVSWSVQDLLPLCQELNIPLVLDWHHHNIVPGTLREGSLDLM 250
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   251 PLIPTIRETWTRKGITQKQHYSESADPTAISGMKRRAH-----SDRVFDF-----PPCDP 300
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              228 ETWORIVATWEDSPLPVKIHLS ----SPLSGEDDPRHHDYINADRFIAFLHEIGADAVD 282
halodurans and genomic sequence comparison with Bacillus subtilis.";
Nucleic Acids Res. 28:437-4331(2000).
EMBL, APO01516; BAB06316.1; -.
CCOMPLETE PROTECTION AND ABODC9291CD1E7B4 CRC64;
                                                                                                                                                                                                                                                                                                        37; Gaps
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          for DNA
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     81 HFMRVSSDLFP---FASHAKYGYTLEFAQSHLEEVGKLANKYNHRLTMHPGQYTQIASPR 137
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 138 EVVVDSAIRDLAYHDEI----LSRMKLNEQLNKDAVLIIHLGGTFEGKKETLDRFRKNY 192
                                                                                                                                                                                                                                                                                                                                                                  19 RGRLGYACLNTILRSMKERVFCSRTCRITTIQR-----DGLESVKQLGTQNVLDLIKLVE 73
                                                                                                                                                                                                                                                                                                                                                                                                                                2 RIQFGYVAMSMELANASP----SKTMTATOFEKIEDHEAGLRKLERIAKTNLHNCLRLLK 57
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           Gaps
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   68 RLYRLSSSLFPMLDLAGDDTGAAVLTHLAPQLLEAGHAFTDAGVRLLMHPEQFIVLNSDR 127
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  21 RLGYACLNTILRSMKERVFCSRTCRITTIQRDGLESVKQLGTQNVLDLIKLVEWNHNFGI 80
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   10 OLGLVCLTVGPEVRFRTVTLSRYRALSPAEREA - - KLLDLYSSNIKTLRGAADYCAAHDI
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  Bacteria; Thermus/Deinococcus group; Deinococcales; Deinococcus.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             Kitayama S., Kikuchi M., Funayama T., Narumi I., Watanabe H.; "Cloning of structural gene of an alternative incision enzyme damage in Delnococcus radiodurans."; Submitted (OCT-1999) to the EMBL/GenBank/DDBJ databases.

EMBL; AB033747; BAA85759.1; -.
                                                                                                                                                                                                                                          Length 322;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  11.6%; Score 229.5; DB 2; Length 305; 27.9%; Pred. No. 1.6e-09;
                                                                                                                                                                                                                                                                                                           Indels
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           Indels
                                                                                                                                  322 AA; 36900 MW; 4B0DC9291CD1E7B4 CRC64;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              305 AA; 33592 MW; B94D333243E2FEA4 CRC64;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           01-MAY-2000 (TrEMBLrel. 13, Created)
01-MAY-2000 (TrEMBLrel. 13, Last sequence update)
01-MAY-2000 (TrEMBLrel. 13, Last annotation update)
                                                                                                                                                                                                                            15.5%; Score 306.5; DB 2; 26.6%; Pred. No. 3e-15; tive 75; Mismatches 123;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       305 AA
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       Mismatches
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       PRT;
```

```
10;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           138 EVVVDSAIRDLAYHDEI----LSRMKLNEQLNKDAVLIIHLGGTFEGKKETLDRFRKNY 192
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     | : | : | : | : | : | : | 149 PEVRESSVRAMSAHARVMDGLGLARTPWN------LLLLHGGKGGRGAE-----LAALI 196
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                252 -LIPTIRETWTRKGIT-----QKQHYSESADPTAISGMKRRAHSDRVFDFPPCDPTMD- 303
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 248 QEDPSVREWVLRARATWQPPEWQVVHLS-----NGIEGPQDRRHSHLIADFPSAYADVPW 302
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           81 HFMRVSSDLFP---FASHAKYGYTLEFAQSHLEEVGKLANKYNHRLTMHPGQYTQIASPR 137
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 89 RLYRLSSSLFPMLDLAGDDTGAAVLTHLAPQLLEAGHAFTDAGVRLLMHPEQFIVLNSDR 148
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          193 QRLSDSVKARLVLENDDVSWSVQDLLPLCQELNIPLVLDWHHNIVPGTLREGSLDLMP- 251
               128 PEVRESSVRAMSAHARVMDGLGLARTPWN-----LLLLLHGGKGGRGAE----LAALI 175
                                                                                                                                         252 -LIPTIRETWTRKGIT-----QKQHYSESADPTAISGMKRRAHSDRVFDFPPCDPTMD- 303
                                                                                                                                                                 Gaps
                                                           193 QRLSDSVKARLVLENDDVSWSVQDLLPLCQELNIPLVLDWHHHNIVPGTLREGSLDLMP- 251
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         21 RLGYACLNTILKSMKERVFCSRTCRITTIQRDGLESVKQLGTQNVLDLIKLVEWNHNFGI 80
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 31 QLGLVCLTVGPEVRFRTVTLSRYRALSPAEREA--KLLDLYSSNIKTLRGAADYCAAHDI 88
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                White O., Eisen J.A., Heidelberg J.F., Hickey E.K., Peterson J.D., Dodson R.J., Haft D.H., Gwinn M.L., Nelson W.C., Richardson D.L., Moffat K.S., Qin H., Jiang L., Pamphile W., Crosby M., Shen M., Vamathevan J.J., Lam P., McDonald L., Utterback T., Zalewski C., Makarova K.S., Aravind L., Daly M.J., Minton K.W., Fleischmann R.D., Ketchum K.A., Nelson K.E., Salzberg S., Smith H.O., Venter J.C.,
                                                                               Bacteria; Thermus/Deinococcus group; Deinococcales; Deinococcus.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 45;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          DB 2; Length 326;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              "Genome sequence of the radioresistant bacterium Deinococcus
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     Query Match 11.6%; Score 229.5; DB 2; Length 3 Best Local Similarity 27.9%; Pred. No. 1.8e-09; Matches 88; Conservative 42; Mismatches 140; Indels
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          326 AA; 35693 MW; C4EA0DOAD2C38988 CRC64;
                                                                                                                                                                                                                                                                                                                                                                                                                                      MAY-2000 (TrEMBLrel. 13, Last sequence update)
MAR-2001 (TrEMBLrel. 16, Last annotation update)
DAMAGE ENDONUCLEASE, PUTATIVE.
-
::
::
                                                                                                                                                                                                                                                                                                                                                                                                    01-MAY-2000 (TrEMBLrel. 13, Created) 01-MAY-2000 (TrEMBLrel. 13, Last sequ 01-MAR-2001 (TrEMBLrel. 16 Tast north
                                                                                                                                                                                                                                                                                                                                                                              PRT;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 MEDLINE=20036896; PubMed=10567266;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      Endonuclease; Complete proteome.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            EMBL; AE002022; AAF11370.1; -.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          Science 286:1571-1577(1999).
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                304 LMIEAKEKEQAVFEL 318
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       Deinococcus radiodurans.
                                                                                                                                                                                                                               304 LMIEAKEKEQAVFEL 318
                                                                                                                                                                                                                                                                      282 IEVEAKGKEEAIAAL 296
                                                                                                                                                                                                                                                                                                                                                                              PRELIMINARY;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           SEQUENCE FROM N.A.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 NCBI_TaxID=1299;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        radiodurans R1.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 Fraser C.M.;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                SEQUENCE
                                                                                                                                                                                                                                                                                                                                                                                                Q9RTE6;
                                                                                                                                                                                                                                                                                                                                                                              Q9RTE6
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               1
                                                                                                                                                                                                                                                                                                                                    RESULT 7
                                                                                                                                                                                                                                                                                                                                                          Q9RTE6
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   QΩ
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  Db
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            ά
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              δλ
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      g
                                                                                                                                                                                   g
                                                                                                                                                                                                                             ōλ
                                                                                                                                                                                                                                                                                                                                                                              δλ
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           δ
                                                                                                   g
                                                                                                                                             δ
```

```
13;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             194 -----QAGL--VKHTG-KSVLD-----EHEDNFALVFAAMGVNM------ETARFF 230
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    106 QSHLEEVGKLANK-----D 142
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         231 KODFEENGSMENVCLFLNLANDPTIERIITPRLALTAAEFLAYOCEKHVLVILTDMSSYA 290
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     291 EALREVSAAREEVPGRRGFPGYMYTDLATIYERAGRVEGRNGSITQI------P 338
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            202 RLVLENDDVSWSVQDLLPLCQELNIPLVLDWHHHNIVPGTLREGSLDLMPLIPTIRETWT 261
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               339 ILTMPNDDITHPIPDLTGXITEGQIYVDRQLHNRQIYP-----PVNVLPSLSRLMKSAI 392
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               134 DFLDIQGQPINPWSRIYPEEMIQTGISAIDVMNSIARGOKIPIFSAAGLPHNEIAAQICR 193
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         46 ITTIQRDGLESVKQLGTQNVLDLIKLVEWNHNFGIHFMRVSSDLFPFASHAKYGYTLEFA 105
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                143 SAIRDL-AYHDEILSRMKLNEQLNKDAVLIIHLGGTFEGKKETLDRFRKNYQRLSDSVKA 201
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        86; Gaps
                                                                                                                                                                                                                                                    Aedes aegypti (Yellowfever mosquito).
Eukaryota: Metazoa; Arthropoda; Tracheata; Hexapoda; Insecta;
Pterygota; Neoptera; Endopterygota; Diptera; Nematocera; Culicoidea;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             7 EMFDCLDKPI-PW-----CSRTCRYACLNTILRSMKERVF------CSRTCR 45
                                                                                                                                                                                                                                                                                                                                                                                       Cociancich S.O., Shahabuddin M.; "Vesicular ATPase over-expressing cells determine the pattern of Plasmodium gallinaceum oocyst distribution in the midguts of Aedes
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 5.6%; Score 111; DB 5; Length 496; 20.1%; Pred. No. 2.3;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 ; Pred. No. 2.3;
47; Mismatches 114; Indels
                                                                                                                                                                                                                                                                                                                                                                                                                                            aegypti and Anopheles gambiae.";
Submitted (SEP-1998) to the EMBL/GenBank/DDBJ databases.
-: SIMILARITY: TO ATPASE ALPHA/BETA CHAINS FAMILY.
EMBL: AF092934: AADD7666.1; -.
Interpro; IPR000194; ATPase_alpha_beta.
Interpro; IPR003255; ATP_synt_B_c.
Pfam; PF00006; ATP_synt_B_c.
Pfam; PF00008; ATP_synt_B_c.i.
PROSITE; PS00152; ATPASE_ALPHA_BETA; 1.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     Hydrogen ion transport; Hydrolase.
SEQUENCE 496 AA; 55216 MW; 8CABC880213F4BEE CRC64;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              01-JUN-1998 (TrEMBLrel. 06, Created)
01-JUN-1998 (TrEMBLrel. 06, Last sequence update)
01-JUN-2001 (TrEMBLrel. 17, Last annotation update)
B SUBURT V-ATPASE (EC 3.6.1.34).
                                                                                                                                                                        01-NOV-1999 (TrEMBLrel. 12, Last sequence update) 01-JUN-2001 (TrEMBLrel. 17, Last annotation update)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          492 AA.
                                                                                                                  496 AA
                                                                                                                                                                                                               VACUOLAR ATPASE B SUBUNIT (EC 3.6.1.34).
                                                                                                                                                       01-NOV-1999 (TrEMBLrel. 12, Created)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          PRT;
                                                                                                                  PRT;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          62; Conservative
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              FRELIMINARY;
303 IEVEAKGKEEALAAL 317
                                                                                                                PRELIMINARY;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     Query Match
Best Local Similarity
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        262 RKGITQKQH 270
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             393 GEGMTRKDH 401
                                                                                                                                                                                                                                                                                                                                                                       SEQUENCE FROM N.A.
                                                                                                                                                                                                                                                                                                                                  NCBI_TaxID=7159;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            044427;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            044427
                                                                                                                                   Q9XYC8;
                                                                                                                  09XYC8
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       S
                                                                           œ
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            Matches
                                                                                                                                                                                                                                     VATB.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     RESULT
                                                                             RESULT
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           044427
                                                                                                09XYC8
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     qq
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        QΥ
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              a
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               δy
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  qq
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         δò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             q
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              δ
                    qq
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   δ
```

```
PRELIMINARY;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     Eukaryota; Metazoa; (
Ascidiidae; Ascidia.
NCBI_TaxID=79730;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         Ascidia sydneiensis
      SEQUENCE FROM N.A.
                                                                                         SEQUENCE FROM N.A.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       SEQUENCE FROM N.A.
                                                                                                        TISSUE=GILL;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 .990960
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  990960
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            RESULT 11
      RP
RRT
RRT
RRT
RRT
RRT
DDR
DDR
DDR
SO
                                                                                                                                                                                                                                                                                                                                                                                                qq
                                                                                                                                                                                                                                                                                                                                                                          Óλ
                                                                                                                                                                                                                                                                                                                                                                                                                           δ
                                                                                                                                                                                                                                                                                                                                                                                                                                                  g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                          δ
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  QQ
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   q
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   QQ
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    qq
                                                                                                                                                                                                                                                                                              13;
                                                                                                                                                                                                                                                                                                                                        130 DFLDIQGQPINPWSRIYPEEMIQTGISAIDVMNSIARGQKIPIFSAAGLPHNEIAAQICR 189
                                                                                                                                                                                                                                                                                                                                                                   46 ITTIQRDGLESVKQLGTQNVLDLIKLVEWNHNFGIHFMRVSSDLFPFASHAKYGYTLEFA 105
                                                                                                                                                                                                                                                                                                                                                                                                                    227 KODFEENGSMENVCLFLNLANDPTIERIITPRLALTAAEFLAYOCEKHVLVILTDMSSYA 286
                                                                                                                                                                                                                                                                                                                                                                                                                                                                    143 SAIRDL-AYHDEILSRMKLNEQLNKDAVLIIHLGGTFEGKKETLDRFRKNYQRLSDSVKA 201
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    202 RLVLENDDVSWSVQDLLPLCQELNIPLVLDWHHHNIVPGTLREGSLDLMPLIPTIRETWT 261
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       Weihrauch D., Lignot J.-H., Towle D.W.;
"Modecular characterization of the B-subunit of a v-type H+-ATPase in
gills of the euryhaline shore crab Carcinus maenas.";
Am. Zoologist 0:0-0(1999).
           Eukaryota; Metazoa; Arthropoda; Tracheata; Hexapoda; Insecta;
Pterygota; Neoptera; Endopterygota; Diptera; Nematocera; Culicoidea;
                                                                                                                                                                                                                                                                                              86; Gaps
                                                                                                                                                                                                                                                                                                                 7 EMFDCLDKPI-PW-----CSRTCR 45
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 01-MAY-2000 (TrEMBLrel. 13, Created)
01-MAY-2000 (TrEMBLrel. 13, Last sequence update)
01-UN-2001 (TrEMBLrel. 17, Last annotation update)
VACUOLAR ATP SYNTHASE SUBUNIT B (EC 3.6.1.35) (VACUOLAR ATP SYNTHASE
                                                                                                                                                                                                                                                                                                                                                                                           -----OAGL--VKHTG-KSVLD-----DHEENFAIVFAAMGVNM------ETARFF
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             EALREVSAAREEVPGRRGFPGYMYTDLATIYERAGRVEGRNGSITQI------P
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              Carcinus maenas (Common shore crab) (Green crab).
Bukaryota: Metazoa; Arthropoda: Crustacea; Malacostraca;
Bumalacostraca; Eucarida; Decapoda; Plecyemata; Brachyura;
Eubrachyura; Portunoidea; Portunidae; Carcinus.
                                                                                                                                                                                                                                                                  5.5%; Score 109; DB 5; Length 492; 19.7%; Pred. No. 3.2;
                                                                                                                                                                                                                                                                                           48; Mismatches 114; Indels
                                                                                                                                                                    PROSITE; PS00152; ATP-Synt_Bc; 1.
PROSITE; PS00152; ATPASE_ALPHA_BETA; 1.
Hydrogen ion transport; Hydrolase.
SEQUENCE 492 AA; 54778 MW; 2EB6CB60DA8FA570 CRC64;
                                                                                                     Culex quinquefasciatus (Southern house mosquito).
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                489 AA.
                                                                         SEQUENCE FROM N.A.
TISSUE=MIDGUT AND MALPIGHIAN TUBULES;
                                                                                                Filippova M., Ross L.S., Gill S.S.;
Insect Mol. Biol. 7:0-0(1998).
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               PRT;
                                                                                                                                                                                                                                                                                           Conservative
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               PRELIMINARY;
                                                                                                                                                                                                                                                                              Similarity
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      262 RKGITQKQH 270
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              389 GEGMTRKDH 397
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         SEQUENCE FROM N.A.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       SUBUNIT B L FORM)
                                                NCBI_TaxID=7176;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              NCBI_TaxID=6759;
                                                                                                                                                                                                                                                                   / Mu
Local Sim
                                                                                                                                                                                                                                                                  Query Match
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           Q9U6A8;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            09U6A8
                                                                                                                                                                                                                                                                                                                                                                                            190
                                                                                                                                                                                                                                                                                        Matches
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       .
[2]
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    RESULT
ò
                                                                                                                                                                                                                                                                                                                                                                                          QQ
                                                                                                                                                                                                                                                                                                                                                                                                                                           g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            q
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      ò
                                                                                                                                                                                                                                                                                                                                                                                                                    ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                     ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              셤
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    ŏ
```

```
12;
Weihrauch D., Lignot J.-H., Towle D.W.;
"Molecular characterization of the B-subunit of a v-type H+-ATPase in gills of the euryhaline shore crab Carcinus maenas.";
Am. Zoologist 39:66A-66A(1999).
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          Chordata; Urochordata; Ascidiacea; Phlebobranchia;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               Gaps
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           118 DQGPSVLAEDFLDIQGQPINPWSRTYPEEMIQTGISSIDVMNSIARGQKIPIFSAAGLPH 177
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      178 NEIAAQICR-----QAGLVKLPSKGVHD------DHKDNFAIVFAAMGVNM----- 217
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              97 KYGYTLEFAQSHLEEVGKLANK-------YNHRLTMHPGQYTQIASPREVVV 141
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    218 ---ETARFFKODFEENGSMENVCLFLNLANDPTIERIITPRLALITAEYLAYOCEKHVLI 274
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            142 ------DSAIRDL-AYHDEILSRMKLNEQLNKDAVLIIHLGGTFEGKKETLDRFFRKNY 192
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  275 ILTDMSSYAEALREVSAAREEVPGRRGFPGYMYTDLATIYERAGRVEGRQGSITQI---- 330
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           193 QRLSDSVKARLVLENDDVSWSVQDLLPPLCQEL.NTPLVLDWHHHN1VPGTLREGSLDLMPL 252
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                331 -----PILTMPNDDITHPIPDLTGYITEGOIYVERQLHNRQIYP-----PINVLPS 376
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            2 DHAP----REMFDCLDKPI-PW------39
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             40 --- CSRTCRITTIORDGLESVKQLGTQNVLDLIKLVEWNHNFGIHFMRVSSDLFPFASHA 96
                                                                                                                                                                                                                                  "Splice variant of V-ATPase subunit B mRNA from gills of shore crab
Carcinus maenas.";
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                TISSUE=BLOOD;
Ucki T., Uyama T., Kanamori K., Michibata H.;
"Isolation of cDNAs encoding A- and B-subunits of the Vacuolar-Type
ATPase from the ascidian Ascidia sydnensis samea.";
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               90;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             Query Match 5.5%; Score 108; DB 5; Length 489; Best Local Similarity 18.2%; Pred. No. 3.8; Matches 58; Conservative 49; Mismatches 121; Indels
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         Mismatches 121; Indels
                                                                                                                                                                                                                                                                                  Submitted (MAR-2000) to the EMBL/GenBank/DDBJ databases.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  Hydrogen ion transport; Hydrolase.
SEQUENCE 489 AA; 54516 MW; 55F4C967654B0648 CRC64;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          01-MAY-1999 (TrEMBLrel. 10, Created)
01-MAY-1999 (TrEMBLrel. 10, Last sequence update)
01-JUN-2001 (TrEMBLrel. 17, Last annotation update)
VACUOLAR-TYPE H+-ATPASE SUBUNIT B (EC 3.6.1.34).
                                                                                                                                                                                                                                                                                                                                                                                                InterPro; IPR000194; ATPBASE_alpha_beta.
InterPro; IPR003255; ATP_synt_B_c.
Pfam; PF00006; ATP-synt_ab; 1.
Pfam; PF01038; ATP-synt_B_c; 1.
PROSITE; PS00152; ATPASE_ALPHA_BETA; 1.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 PRT;
                                                                                                                                                                                                                                                                                                                                             EMBL; AF189779; AAF08281.1; -. EMBL; AF247971; AAF67183.1; -.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      253 IPTIRETWTRKGITQKQH 270
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  377 LSRLMKSAIGEGMTRKDH 394
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            Conservative
                                                                                                                                                                                                    Weihrauch D., Towle D.W.;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  samea
```

```
338 ----PILTMPNDDITHPIPDLTGYITEGQIYIDRQLHNRQIYP------PINVLPSLSRL 387
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    257 RETWTRKGITQKQH-----YSESADPTAISGMKRRAHSDRVFDFPPCDPTMDLMI--- 306
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      307 -----EAKEKEQAVF----ELCRRYELQNPPCPLEIMGPEYDQTRDGYYPP 348
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 440 FLGKFEKTFIAQGPYENRSVFDSLNIGWELLRIF------PKEMLKRIPRNVINEYYP- 491
                                                                                                                                                                                                                                                                                                                                                                                                                                                                  192 CR-----OGGL---VKLPDKDVMD-----SHEDNFAIVF------AAMGVNME 225
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             104 ---FAQSHLEEVGKLANK-----141
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                142 ----DSAIRDL-AYHDEILSRMKLNEQLNKDAVLIIHLGGTFEGKKETLDRFRKNYQRLS 196
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      286 MSSYAEALREVSAAREEVPGRRGFPGYMYTDLATIYERAGRVNGRNGSITQI----- 337
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   197 DSVKARLVLENDDVSWSVQDLLPLCQELNIPLVLDWHHHNIVPGTLREGSLDLMPLIPTI 256
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               388 MKSAIGEGMTRKDHSDVSNQLYANYAIGKDVQAMKAVVGEEAL-----TADDMLYLE 439
                                                                                                                                                                                                                                                                                                                                                                                                                         44 CRITTIQRDGLESVKQLGTQNVLDLIKLVEWNHNFGIHFMRVSSDLFPFASHAKYGYTLE 103
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     226 AARFFKSDFEQHGSMDNVCLFLNLANDPTIERIITPRIALTTAEFLAYQCEKHVLVILTD 285
                                                                                                                                                                                                                                                                                                                                                                                   132 PEDYLDIQGQPINPKSRIYPEEMIQTGISAIDTMNSIARGQKIPIFSANGLPHNEIAAQI 191
                                                                                                                                                                                                                                                                              ; Pred. No. 6.7; 69; Mismatches 151; Indels 144; Gaps
                                                                                                                                                                                                                                                                                                                                        5 PREMFDCLDKPIPWRGRL------GYACLNTILRSMKERVF------CSRT 43
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             Eukaryota; Alveolata; Ciliophora; Oligohymenophorea; Peniculida;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     Proissard M., Keller A.M., Cohen J.;
"Nd9p, a novel protein with armadillo-like repeats involved in
exocytosis: physiological studies using allelic mutants in
Paramecium.";
                                                                                                                                                                                                                                                          DB 5; Length 509;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        Submitted (SEP-2000) to the EMBL/GenBank/DDBJ databases. EMBL, AJ293945; CAC12829.1; -- SEQUENCE 881 AA; 104104 MW; 8E27791A74CDD4E6 CRC64;
Submitted (JUL-1998) to the EMBL/GenBank/DDBJ databases.
-!- SIMILARITY: TO ATPASE ALPHA/BETA CHAINS FAMILY.
EMBL; AB016484; BAA36692.1; -.
                                                                                                     Pfam; PF00006; ATP-synt_ab; 1.
Pfam; PF01038; ATP-synt_ab; 1.
PROSITE; PS00152; ATPASE_ALPHA_BETA; 1.
Hydrogen ion transport; Hydrolase.
SEQUENCE 509 AA; 56675 MW; 185BED824376451B CRC64;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  01-WAR-2001 (TrEMBLrel. 16, Created)
01-WAR-2001 (TrEMBLrel. 16, Last sequence update)
01-WAR-2001 (TrEMBLrel. 16, Last annotation update)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           881 AA
                                                                                                                                                                                                                                                     Query Match 5.3%; Score 105;
Best Local Similarity 17.8%; Pred. No. 6
                                                               InterPro; IPR000194; ATPase_alpha_beta.
InterPro; IPR003255; ATP_synt_B_c.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              PRT;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             349 GAEKRLTARKRRSRKEEVEEDEK 371
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    492 -----RKKVPSAQDKQEDQK 506
                                                                                                                                                                                                                                                                                                  79; Conservative
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              PRELIMINARY;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                Paramecium tetraurelia.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     SEQUENCE FROM N.A.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           NCBI_TaxID=5888;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   ND9 PROTEIN.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           Paramecium
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              Q9GRI1
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   RESULT 12
                                                                                                                                                                                                                                                                                                    Matches
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         Q9GRI1
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    Ω
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                qq
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    QQ
                                                                                                                                                                                                                                                                                                                                                                                                                              οy
                                                                                                                                                                                                                                                                                                                                                                                                                                                                    qq
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               δ
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                Qγ
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   οy
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       ŏ
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               임
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             δ
    CCC
DR
DR
DR
DR
SQ
SQ
SQ
```

```
11;
                                           18;
                                                                                                                                                                                                                                                                                                                                                                 -----TY 308
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               170 MNSIARGOKIPIFSAAGLPHNEIAAQICR-----QAGL--VKK--SKDVMDXS----DD 215
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         ---YNH 121
                                                                                                                                                                                            91 HTLERRDLIKFIQGGWKTDYMYKFLEVRELPTYKGKFRDYNYQVTMLKKPQVELMENQEM 150
                                                                                                                                                                                                                                                                                           151 YKFNMHVLNGYNIFFLNSY-----QNTKKGLYRNGQNKSQFTLQVSETNELEVLEHMHNH 205
                                                                                                                                                                                                                                                                                                                                  146 RDLAYHDEILSRMKENEQLNKDAVLIIHLGGTFEGKKETLDRFRKNYQRLSDSVKARLVL 205
                                                                                                                                                                                                                                                                                                                                                                                                                 206 ENDDVSWSVQDLLPLCQELNIPLVLDWHHHNIVPGTLREGSLDLMPLIPTIRETWTRKGI 265
                                                                                                                                                                                                                                                                                                                                                                                                                                                         266 TQKQHYSESADPTAISGMKRRAHSDRVFDFPPCDPTMDLMIEAKEKEQAVFELCRRYELQ 325
                                                                                                                                                                                                                                                   90 FPFASHAKYGYTLEFAQSHLEEVGKLANKYNHRLTMHPGQYT-QIASPREVVVDSAI--- 145
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          27 LNTILRSMKERVF-----CSRTCRITTIQRDGLESVKQLGTQNVLDLIKLVEWNH 76
                                              Gaps
                                                                                                                          44 KEGFDILE-----AKVSFQCIDRLIVS-----SQDIYLFQIHFNGKRPKNVPLKMNI 90
                                                                                   6 REMFDCLDKPIPWRGRLGYACLNTILRSMKERVFCSRTCRITTIQRDG-----LESVKQL 60
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         "Cloning and molecular characterization of the trout (Oncorhynchus mykiss) vacuolar H+-ATPase B subunit.";
Submitted (APR-1999) to the EMBL/GenBank/DDBJ databases.
--- SIMILARITY: TO ATPASE ALPHA/BETA CHAINS FAMILY.
EMBL; AF140022; AAD33861.1;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      01-NOV-1999 (TrEMBLrel. 12, Created)
01-NOV-1999 (TrEMBLrel. 12, Last sequence update)
01-JUN-2001 (TrEMBLrel. 17, Last sequence update)
01-JUN-2001 (TrEMBLrel. 17, Last annotation update)
V-TYPE ATPASE B SUBBUIT (EC 3.6.1.34)
V-TYPE ATPASE B SUBBUIT (EC 3.6.1.34)
Concortynchus mykiss (Rainbow trout) (Salmo gairdneri).
Eukaryotas Metasoas Chordata; Craniata; Vertebrata; Euteleostomi;
Actinopterygii; Neopterygii; Teleostei: Buteleostei;
Protacanthopterygii; Salmoniformes; Salmonidae; Oncorbynchus.
                                         61; Mismatches 126; Indels 144;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    5.2%; Score 103; DB 13; Length 502; 19.4%; Pred. No. 9.3;
5.3%; Score 104; DB 5; Length 881; 18.5%; Pred. No. 17;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          273 TTKQAIEME---YAVIMMRRKYHPPYLENF-----VDIVLTF----
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  326 NPPCPLETMGPEYDOTRDGYYPPGAEKRLTARKRRSRKEEVEEDEK 371
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             309 DPKCKVEDNNPEADQD------QQELEEQLQESENEEDGEEEK 345
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             77 NFGIHFMRVSSDLFPFASHAKYGYTLEFAQSHLEEVGKLANK-----
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               OFC8CBE972D22BA0 CRC64;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  :66
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     502 AA
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    49; Mismatches
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               InterPro; irrocally pfam; PF00006; ATP-synt_ab; 1. Pfam; PF01038; ATP-synt_B_c; 1. PROSITE; PS00152; ATPASE_ALPHA_BETA; 1. "...aranan ion transport; Hydrolase."
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    Interpro; IPR000194; ATPase_alpha_beta.
Interpro; IPR003255; ATP_synt_B_c.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                Perry S.F., Beyers M.L., Johnson D.A.;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                PRT;
                                                                                                                                                                      61 GTQNVLDLIKLVE--WNHNFGIHFMRV---
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             55738 MW;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    Conservative
                          Best Local Similarity 18.59
Matches 75; Conservative
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  PRELIMINARY;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         Hydrogen ion transport;
SEQUENCE 502 AA; 557
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              Local Similarity
nes 54; Conserv
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      SEQUENCE FROM N.A.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          NCBI_TaxID=8022
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              TISSUE=GILL;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               Ouery Match
        Query Match
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           09W6M4;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     09W6M4
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     Best Loca
Matches
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             RESULT 13
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   29W6M4
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           δ
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        ga
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               δÀ
                                                                                                                                                                                                                                                                                                                                                                              gg
                                                                                                                                                                                                                                                                                                                                                                                                                                                             Db
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         ŏλ
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              pp
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        Óγ
                                                                                                                                                                                                              Dp
                                                                                                                                                                                                                                                                                             qq
                                                                                                                                                                                                                                                                                                                                       Óγ
                                                                                                                                                                                                                                                                                                                                                                                                                        ŏ
                                                                                                                               q
                                                                                                                                                                         ŏ
                                                                                                                                                                                                                                                          Qγ
```

```
16;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 232 WHHHNIVPGTLREGSLDLMPLIPTIRETWTRKGITQKQH-----YSESADPTAISGMK 284
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               253 PRIALTTAEYLAYECGKHVLVILTDMSSYADALREVSAAREEVPGRRGYPGYMYTDLATI 312
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            313 YERAGRIEGRKGSITQI------PILTMPNDDITHPTPDLTGYITEGQIYIDRQ 360
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     361 LHNRQIYP-----PINVLPSLSRLMKSAIGEGMTRRDHSDVSNQLYANYAIGKDVQAMK 414
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          285 RRAHSDRVFDFPPCDPTMDLM-IEAKEKEQAVFELCRRYELQN-------PPCPL 331
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                :|:| | | | :| : | : | | : | | : | | | : | | | | : | | | | : | | : | | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | 
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               76 HNFGIHFMRVSSDLFPFASHAKYGYTLEFAQSHLEEVG------KLANK-----YN 120
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             172 IHLGGTFEGKKETLDRFRKNYQRLSDSVKARLVLENDDVSWSVQDLLPLCQELNIPLVLD 231
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        203 ENFAIVFAAMGVNM------ETAOFFKRDFEENGSMERVTLFLNLANDPIIERIIT 252
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        121 HRLTMHPGQYTQIASPREVVV-----DSAIRDL-AYHDEILSRMKLNEQLNKDAVLI 171
                                                                                                                                                                                      266 RLALTSAEYLAYQCEKHVLVILTDMSSYAEALREVSAAREEVPGRRGFPGYMYTDLATIY 325
                                                                                                                                                                                                                                                                               173 HLGGTFEGKKETLDRFRKNYQRLSDSVKARLVLENDDVSWSVQDLLPLCQELNIPLVLDW 232
Gaps
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           27 LNTILRSMKERVF------CSRTCRITTIQRDGLESVKQL-GTQNVLDLIKLVEWN 75
                                                                                                          122 RLTMHPGQYTQIASPREVVV-----DSAIRDL-AYHDEILSRMKLNEQLNKDAVLII 172
                                                                                                                                                                                                                                                                                                                                                               326 ERAGRVEGRNGSITQI------PILTMPNDDITHPIPDLTGYITEGOVYVDRQL 373
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta; Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; Ehrhartoideae; Oryzeae; Oryza.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  Query Match
5.2%; Score 102.5; DB 10; Length 488;
Best Local Similarity 19.5%; Pred. No. 9.8;
Matches 73; Conservative 63; Mismatches 138; Indels 101;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      STRAIN=CV. NIPPONBARE;
Sasaki T., Matsumoto T., Yamamoto K.;
"Oryza sativa nipponbare(GA3) genomic DNA, chromosome 1, PAC
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      Submitted (NOV-2000) to the EMBL/GenBank/DDBJ databases.
EMBL; AP002901; BAB39419.1; --
SEQUENCE 488 AA; 54279 MW; 99C6821732198EFI CRC64;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       01-JUN-2001 (TrEMBLrel. 17, Created)
01-JUN-2001 (TrEMBLrel. 17, Last sequence update)
01-JUN-2001 (TrEMBLrel. 17, Last annotation update)
PUTATIVE H+-TRANSPORTING ATP SYNTHASE.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       374 HNROIYP-----PINVLPSLSRLMKSAIGEGMTRKDH 405
                                                                                                                                                                                                                                                                                                                                                                                                                                                            233 HHHNIVPGTLREGSLDLMPLIPTIRETWTRKGITQKQH 270
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             PRT;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   EIMGPEYDQTRDGYY 346
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            ELLHRIPAKTLDOYY 482
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        PRELIMINARY;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  Oryza sativa (Rice)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      SEQUENCE FROM N.A.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  clone: P0456F08."
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      NCBI_TaxID=4530;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             Q9ASE0
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              D9ASE0
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            468
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 a
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 á
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             q
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 ŏ
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      g
                                                                                                                                                                                                                                                                                                ò
```

```
14;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   336 ------PILTMPNDDITHPIPDLTGYITEQ1YVDRQLHNRQIYP-----PINVLPS 381
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              97 KYGYTLEFAQSHLEEVGKLANK------YNHRLTMHPGQYTQIASPREVVV 141
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            193 QRLSDSVKARLVLENDDVSWSVQDLLPLCQELNIPLVLDWHHHNIVPGTLREGSLDLMPL 252
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            184 NEIAAQICR-----QAGL--VKK--SKDVMDYSE-----ENFAIVFAAMGVNM----- 222
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        223 ---ETARFFKSDFEENGSMDNVCLFLNLANDPTIERITTPRLALTTAEFLAYQCEKHVLV 279
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     -----DSAIRDL-AYHDEILSRMKLNEQLNKDAVLIIHLGGTFEGKKETLDRFRKNY 192
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                124 DRGPAVLAEDFLDIMGQPINPQCRIYPEEMIQTGISAIDGMNSIARGOKIPIFSAAGLPH 183
                                                                                                                                                                  Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Archosauria; Aves; Neognathae; Galliformes; Phasianidae; Phasianinae;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              Gaps
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          40 ---CSRTCRITTIQRDGLESVKQLGTQNVLDLIKLVEWNHNFGIHFMRVSSDLFPFASHA 96
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 280 ILTDMSSYAEALREVSAAREEVPGRRGFPGYMYTDLATIYERAGRVEGRNGSITQI----
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                Query Match
5.1%; Score 101.5; DB 13; Length 496;
Best Local Similarity 19.2%; Pred. No. 12;
Matches 61; Conservative 55; Mismatches 111; Indels 91;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       2 DHAP----REMFDCLDKPIPWRGRL------GYAC---LNTILRSMKERVF
                                    01-0CT-2000 (TrEMBLrel. 15, Created)
01-0CT-2000 (TrEMBLrel. 15, Last sequence update)
01-JUN-2001 (TrEMBLrel. 17, Last annotation update)
VACUOLAR H-ATPASE B SUBUNIT OSTEOCLAST ISOZYWE (FRAGMENT).
                                                                                                                                                                                                                                                                                                                                       Submitted (JUN-1996) to the EMBL/Genbank/DDBJ databases.
-! SIMILARITY: TO APPASE ALPHA/BETA CHAINS FAMILY.
EMBL; U61724; AAF73735.1; -.
InterPro; IPR000194; ATPase_alpha_beta.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  NÔN_TER 1 1 1
SEQUENCE 496 AA; 55049 MW; 1CC5D9FDA68BA7BA CRC64;
                                                                                                                                                                                                                                                                                                         Andersen P., Lundberg L.G., Stone D.K.;
"Chicken vacuolar H-ATPase B subunit isozyme.";
496 AA.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               Search completed: January 15, 2002, 14:06:15 Job time: 954 sec
                                                                                                                                                                                                                                                                                                                                                                                                                                  Interpro; IPR003255; ATP_synt_B_c.
Pfam; PF00006; ATP_synt_B_c; 1.
Pfam; PF01038; ATP_synt_B_c; 1.
PROSITE: PS00122; ATPASE_ALPHA_BETA; 1.
Hydrogen ion transport; Hydrolase.
PRT;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      253 IPTIRETWIRKGITOKOH 270
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               382 LSRLMKSAIGECMTRKDH 399
    PRELIMINARY;
                                                                                                                                                       Gallus gallus (Chicken).
                                                                                                                                                                                                                                                                                       SEQUENCE FROM N.A
                                                                                                                                                                                                                                         NCBI_TaxID=9031;
                         Q918A2;
  Q918A2
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              142
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          οp
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 QQ
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              ΟŅ
                         δ
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     ó
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       ŏ
```

.

```
586
 275
6.3
```

GenCore version 4.5 Copyright (c) 1993 - 2000 Compugen Ltd.

OM protein – protein search, using sw model

January 15, 2002, 13:50:09 Run on:

; Search time 120.01 Seconds
(without alignments)
357.991 Million cell updates/sec

Title:

Perfect score:

US-09-724-296-38 3055 1 MGTTGLESL.SLGDRGAAPTV.......HKLNKYHDWLWENHEKLSLS 580 Sequence:

Gapop 10.0 , Gapext 0.5 **BLOSUM62** Scoring table:

522463 seqs, 74073290 residues Searched:

522463 Total number of hits satisfying chosen parameters:

Minimum DB seq length: 0 Maximum DB seq length: 200000000

Post-processing: Minimum Match 0% Maximum Match 100% Listing first 45 summaries

Database :

A_Geneseq_1101:* 1: /SIDS8/qcqdata

/SIDS6/gcgdata/geneseq/geneseqp/AA1980.DAT:*/SIDS8/gcgdata/geneseq/geneseqp/AA1981.DAT:*/SIDS8/gcgdata/geneseq/geneseqp/AA1982.DAT:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

		Description	Human MED1 endonuc	Human delta228-UV	Human ovarian tumo	Secreted protein e	Human secreted pro	Arabidopsis thalia	Arabidopsis thalia	Arabidopsis thalia	BRCA2 cancer susce	Human breast cance	Human breast and o
		ID	AAW74473	AAY44504	AAY76548	AAW88701	AAG02051	AAG16859	AAG16858	AAG16857	AAW25049	. AAW19211	AAW23287
		DB	20	21	20	20	21	21	21	21	18	18	18
	Query	Length	580	265	439	257	219	326	327	352	1644	3418	3418
ф	Query	Match	100.0	8.96	71.8	40.3	6.4	4.9	4.9	4.9	4.7	4.7	4.7
		Score	3055	2958.5	2194.5	1230	194	148.5	148.5	148.5	145	145	145
	Result	S.	1	7	m	4	S	9	7	8	6	10	11

3 120.5 3.9 2665 22 AAM02259 Peptide # 4 120.5 3.9 3266 21 AAB42491 Human ORF 5 119 3.9 646 21 AAB18188 Plasmodiu	1 120.5 3.9 2665 22 AAM14533 Peptide #99 2 120.5 3.9 2665 22 AAM26950 Peptide #93 3 120.5 3.9 2665 22 AAM3259 Peptide #94
RESULT 1 AAW74473 ID AAW74473 XX AAW74473; XX AAW74473; XX AAW74473; XX XX ABOUNClease, MEDI; human; methyl-CpG binding endonuclease-1; Endonuclease; MEDI; human; methyl-CpG binding endonuclease-1; XX XX XX Endonuclease; MEDI; human; methyl-CpG binding andonuclease-1; XX W DNA fidelity; Salamanipulation; Salaesase; spinocerebellar ataxia; XX XX XX Homo sapiens. YA XX XX XX XX YOUR O4-FEB-1999. XX XX XX XX XX XX YOUR O653936. XX XX XX XX YOUR OF SEB-1997; 97US-0053936. XX XX XX XX XX XX XX XX XX	44 120.5 3.9 3266 21 AAB12491 Human ORFX AAW74473 standard; Protein; 580 AA. AAW74473 standard; Protein; 580 AA. AAW74473; 19-MAY-1999 (first entry) Human MEDI endonuclease protein sequence. Endonuclease; MEDI; human; methyl-CpG binding endonuclease-1; Bond fidelity; human; methyl-CpG binding endonuclease-1; Endonuclease; MEDI; human; methyl-CpG binding endonuclease-1; Human MEDI endonuclease protein sequence. Endonuclease; MEDI; human; methyl-CpG binding endonuclease-1; Human MEDI endonuclease protein sequence. Endonuclease; MEDI; human; methyl-CpG binding endonuclease-1; Bondonuclease; MEDI; human; methyl-CpG binding endonuclease-1; Bondonuclease; MEDI; human; methyl-CpG binding endonuclease-1; Endonuclease; MEDI; human; methyl-CpG binding endonuclease-1; Bondonuclease; friplet repeat expansion disorder. ANW ON
ALIGNMENTS 4473 AAW74473; 19-MAY-1999 (first entry) Human MED1 endonuclease protein sequence. Endonuclease; MED1; human; methyl-cpG binding endonuclease-1; DNA fidelity; DNA manipulation; cancer; fragile X syndrome; therapy myotonic dystrophy; Huntington's disease; spinocerebellar ataxia; Homo sapiens. WO9904626-A1. (Ln	44 120.5 3.9 3266 21 AAB12491 Human ORFX Plasmodium ORFX 119 3.9 646 21 AAB12491 Human ORFX Plasmodium AAW74473 Standard; Protein; 580 AA. AAW74473; 19-MAY-1999 (first entry) Human MEDI endonuclease protein sequence. Endonuclease; MEDI; human; methyl-CpG binding endonuclease-1; DNA fidelity; DNA manipulation; cancer; fragile X syndrome; therap myotonic dystrophy; Huntington's disease; spinocerebellar ataxia; Kennedy's disease; triplet repeat expansion disorder. WO9904626-A1. O4-FEB-1999; 28-JUU-1998; 98WO-US15828.
ALIGNMENTS LT 1 AAW74473 AAW74473; 19-MAY-1999 (first entry) Human MED1 endonuclease protein sequence. Endonuclease; MED1; human; methyl-CpG binding endonuclease-1; DNA fidelity; DNA manipulation; cancer; fragile X syndrome; therapy myotonic dystrophy; Huntington's disease; spinocerebellar ataxia; Kennedy's disease; triplet repeat expansion disorder.	44 120.5 3.9 3266 21 AAB18188 Human ORFX Plasmodium AAW74473 AAW74473; Endonuclease, MED1; human; methyl-CpG binding endonuclease-1; DNA fidelity; DNA manpulation; cancer; fragile x syndrome; therap myotonic dystrophy; Huntington's disease; spinocerebellar ataxia; Kennedy's disease; triplet repeat expansion disorder.
LT 1 4473 AAW74473 standard; Protein; 580 AA. AAW74473; 19-MAY-1999 (first entry) Human MED1 endonuclease protein sequence	44 120.5 3.9 3266 21 AAB12491 Human ORFX 45 119 3.9 646 21 AAB18188 Plasmodium ALIGNMENTS ALIGNMENTS ALIGNMENTS ALIGNMENTS ALIGNMENTS AAW74473 standard; Protein; 580 AA. 19-MAY-1999 (first entry) Human MED1 endonuclease protein sequence.
ALIGNMENT LT 1 4473 AAW74473 standard; Protein; 580 AA.	44 120.5 3.9 3266 21 AAB42491 Human ORFX 45 119 3.9 646 21 AAB18188 Plasmodium ALIGNMENTS LT 1 44.73 AAW74473 standard; Protein; 580 AA.
ALIGNMENT LT 1 4473 AAW74473 standard; Protein; 580 AA.	44 120.5 3.9 3266 21 AAB12491 Human ORFX 45 119 3.9 646 21 AAB18188 Plasmodium ALIGNMENTS LT 1 4473 AAW74473 standard; Protein; 580 AA.
	4 120.5 3.9 3266 21 AAR42491 Human ORFX 5 119 3.9 646 21 AAB18188 Plasmodium ALIGNMENTS
	4 120.5 3.9 3266 21 AAB42491 Human ORFX 5 119 3.9 646 21 AAB18188 Plasmodium
1 120.5 3.9 2665 22 AAM14533 Peptide #96 2 120.5 3.9 2665 22 AAM26950 Peptide #96	100: 1 - 1 - 1 - 1 - 1 - 1 - 1 - 1 - 1 - 1
8 121.5 4.0 1105 21 AAG42349 Arabidopsis 9 121.5 4.0 1189 21 AAG42348 Arabidopsis 0 120.5 3.9 2343 20 AAM80899 Canine factor 1 120.5 3.9 2665 22 AAM14533 Peptide #987 2 120.5 3.9 2665 22 AAM26950 Peptide #987	8 121.5 4.0 1105 21 AAG42349 Arabidopsis 9 121.5 4.0 1189 21 AAG42348 Arabidopsis 0 120.5 3.9 2343 20 AAW80989 Canine facto
5 122.5 4.0 995 22 AAM76623 Peptis 6 122.5 4.0 1202 22 AAM77663 Peptis 7 121.5 4.0 1024 21 AAG42350 Arabis 8 121.5 4.0 1105 21 AAG42349 Arabis 9 121.5 4.0 1189 21 AAG42348 Arabis 0 120.5 3.9 2665 22 AAM14533 Peptis 1 120.5 3.9 2665 22 AAM26950 Peptis	5 122.5 4.0 995 22 AAN06223 Peptis 6 122.5 4.0 1202 22 AAN17663 Peptis 7 121.5 4.0 1204 21 AAG42350 Arabis 8 121.5 4.0 1105 21 AAG42349 Arabis 9 121.5 4.0 1189 21 AAG42348 Arabis 0 120.5 3.9 2343 20 AAN80989 Canin
126 4.1 263 19 AAN74980 Human 4 123 4.0 4.1 263 19 AAN74980 Mouse 122.5 4.0 995 22 AAN7663 Peptii 5 122.5 4.0 1202 22 AAN7763 Peptii 7 121.5 4.0 1024 21 AAG42350 Arabis 8 121.5 4.0 1105 21 AAG42349 Arabis 9 121.5 4.0 1189 21 AAG42349 Arabis 120.5 3.9 2645 22 AAN14533 Peptii 120.5 3.9 2665 22 AAN86950 Peptii	126 4.1 263 19 AAW74980 Human 4 123 4.0 414 20 AAV14199 Mouse 5 122.5 4.0 995 22 AAM0523 Pepti, 6 122.5 4.0 1202 22 AAM77663 Pepti, 7 121.5 4.0 1105 21 AAG42349 Arabi, 8 121.5 4.0 1189 21 AAG42349 Arabi, 9 120.5 3.9 2343 20 AAW80989 Canin
120 4.3 1743 19 AAM98879 H. py 120 127.5 4.2 281 20 AAV73829 H. mman 2 127.5 4.2 281 20 AAV748439 H. mman 126 4.1 263 19 AAV448439 H. mman 126 4.1 263 19 AAV74199 H. mman 12.5 4.0 120.2 AAM06223 Peptis 6 122.5 4.0 120.2 2 AAM17663 Peptis 7 121.5 4.0 1105 21 AAG42349 Arabis 9 121.5 4.0 1105 21 AAG42349 Arabis 120.5 3.9 2665 22 AAM14533 Peptis 120.5 3.9 2665 22 AAM14533 Peptis 120.5 3.9 2665 22 AAM14533 Peptis 120.5 3.9 2665 22 AAM16950	130 4.3 1743 19 AAM98879 H. py 120.5 3.9 2.0 AAV73829 Human 127.5 4.2 281 20 AAV73829 Human 127.5 4.2 281 20 AAV74843 Human 126 4.1 263 19 AAW74980 Human 128 4.0 263 19 AAW74980 Human 122.5 4.0 120.2 2 AAW17663 Pepti. Pepti. 121.5 4.0 1024 21 AAG42349 Arabi. 9 121.5 4.0 1189 21 AAG42349 Arabi. 0 120.5 3.9 2343 20 AAW80989 Canin.
8 130 4.3 649 20 AAM90005 Expre 130 4.3 649 20 AAM90005 Expre 130 4.3 867 19 AAW988564 H. py 0 130 4.3 1743 19 AAW98879 H.man 127.5 4.2 281 20 AAY73829 H.man 127.5 4.2 281 20 AAY74980 H.man 128 4.0 414 20 AAY14199 Mouse 5 122.5 4.0 995 22 AAW77663 Peptis 6 122.5 4.0 1202 22 AAW77663 Peptis 6 121.5 4.0 1103 21 AAG42349 Arabis 9 121.5 4.0 1105 21 AAG42349 Arabis 0 120.5 3.9 2665 22 AAW7653 Peptis 120.5 3.9 2665 22 AAW7653 Peptis 120.5 3.9 2665 22 AAW7653 Peptis 120.5 3.9 2665 22 AAW76950	8 130 4.3 649 20 AAM90005 Expre 130 4.3 867 19 AAM98564 H. PY 0 130 4.3 1743 19 AAW98564 H. PY 120 4.2 281 20 AAV73829 Human 2 127.5 4.2 281 20 AAV73829 Human 1 26 4.1 263 19 AAV74980 Human 4 123 4.0 414 20 AAV14199 Mouse 5 122.5 4.0 1202 22 AAM17663 Peptid 7 121.5 4.0 1024 21 AAG42349 Arabi. 9 120.5 4.0 1189 21 AAG42349 Arabi. 9 120.5 3.9 2343 20 AAW80989 Canin
6 130.5 4.3 411 22 AAC64314 Human 130.5 4.3 411 22 AAC64814 Heart 130 4.3 649 20 AAW90005 Expres 130 4.3 649 20 AAW90005 Expres 130 4.3 867 19 AAW98864 Hr. py 0 130 4.3 1867 19 AAW98879 Hr. py 0 127.5 4.2 281 20 AAV73829 Human 2 127.5 4.2 281 20 AAV748439 Human 123 4.0 995 22 AAW14199 Mouse 122.5 4.0 120.2 AAW1763 Pepti 121.5 4.0 1102 22 AAW1763 Pepti 121.5 4.0 1102 21 AAC42349 Arabi 121.5 4.0 1108 21 AAC42349 Arabi 121.5 4.0 1108 21 AAC42349 Arabi 120.5 3.9 2665 22 AAW16989 Canin 120.5 3.9 2665 22 AAW16950 Pepti 120.5 3.9 2665 22 AAW16950	6 130.5 4.3 411 22 AAG64314 Human 130.5 4.3 411 22 AAG64314 Heart 130 4.3 649 20 AAG90005 Expre 9 130 4.3 649 20 AAW99005 Expre 130 4.3 867 19 AAW99879 H. Py 0 127.5 4.2 281 20 AAY73829 Human 127.5 4.2 281 20 AAY48439 Human 126 4.1 263 19 AAW74980 Human 3 126 4.1 263 19 AAW74980 Human 125 4.0 1202 2 AAW1763 Pepti 122.5 4.0 1202 22 AAW1763 Pepti 121.5 4.0 1105 21 AAG42349 Arabi 9 121.5 4.0 1105 21 AAG42349 Arabi 9 120.5 3.9 2343 20 AAW80989 Canin
4 11 22 AAN14197 Human 1 130.5 4.3 411 22 AAA14197 Human 1 130.5 4.3 411 22 AAG64314 Human 1 130.5 4.3 411 22 AAG64314 Human 1 130.5 4.3 411 22 AAG64314 Heart 1 130 4.3 649 20 AAW99005 Expre 9 130 4.3 867 19 AAW998679 H. Py 0 130 4.3 1743 19 AAW99879 H. Py 1 127.5 4.2 281 20 AAY73829 Human 1 127.5 4.2 281 20 AAY73829 Human 3 126 4.1 263 19 AAY74849 Mouse 1 127.5 4.0 995 22 AAW17663 Pepti 1 120.5 4.0 1105 21 AAG42349 AAA1459 AAA1459 AAA1459 AAA1459 AAA1453 AAA1453 Pepti 1 120.5 3.9 2665 22 AAW66989 Canin 1 120.5 3.9 2665 22 AAW66950 Pepti 1 120.5 3.9 2665 22 AAW66950 Pepti 1 120.5 3.9 2665 22 AAW66950 Pepti 1 120.5 3.9 2665 22 AAW66950	4 11 20 AAN14197 Human 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1
2 131.5 4.3 291 20 AAV14198 Human 131 4.3 754 19 AAW33811 Tat s Tat s 130.5 4.3 411 22 AAB99915 Human 5 130.5 4.3 411 22 AAB99915 Human 6 130.5 4.3 411 22 AAB99915 Human 122 AAB99915 Human 130 4.3 649 20 AAW90005 Expres 130 4.3 649 20 AAW90805 Human 127.5 4.2 281 20 AAV3829 Human 127.5 4.2 281 20 AAV4849 Human 127.5 4.2 281 20 AAV4849 Human 127.5 4.2 281 20 AAV4849 Human 127.5 4.0 281 20 AAV4849 Human 127.5 4.0 120.2 AAW14199 Pepti 121.5 4.0 1105 21 AAW42349 AAW1423 Human 121.5 4.0 1105 21 AAW42349 AAW1423 AAW1429 AAW1429 AAW1423 Human 121.5 4.0 1105 21 AAW42349 AAW1423 AAW1423 Pepti 120.5 3.9 2665 22 AAW1453 Pepti 120.5 3.9 2665 22 AAW1453 Pepti 120.5 3.9 2665 22 AAW3650	2 131.5 4.3 291 20 AAV14198 Human 131.5 4.3 754 19 AAV33811 Tat s Human 5 130.5 4.3 411 22 AAS99915 Human 5 130.5 4.3 411 22 AAS99915 Human 6 130.5 4.3 411 22 AAS99915 Human 6 130.5 4.3 411 22 AAS99915 Human 130.5 4.3 411 22 AAS6484 Heart Heart 130 4.3 867 19 AAV98864 Heart Heart 130 4.3 1743 19 AAV98879 Human 127.5 4.2 281 20 AAV73829 Human 2 127.5 4.2 281 20 AAV74800 Human 12.6 4.1 263 19 AAV34980 Human 12.5 4.0 120.2 22 AAV3766.3 Peptit 6 122.5 4.0 120.2 22 AAV3766.3 Peptit 7 121.5 4.0 1024 21 AAG42349 Arabi 8 121.5 4.0 1189 21 AAG42349 Arabi 8 120.5 3.9 2343 20 AAV80989 Canin
137 4.5 560 21 AAB19182 Lipid Mouse 131.5 4.3 291 20 AAY141200 Mouse 131.5 4.3 291 20 AAY141200 Mouse 131.5 4.3 291 20 AAY14120 Mouse 131.5 4.3 291 20 AAY14199 Human 130.5 4.3 411 22 AA809915 Human 130.5 4.3 411 22 AA664314 Heart Human 130.5 4.3 411 22 AA664314 Heart Human 130 4.3 867 19 AAW98879 Human 127.5 4.2 281 20 AAY73829 Human 127.5 4.2 281 20 AAY73829 Human 127.5 4.0 281 20 AAY14199 Mouse 127.5 4.0 1202 22 AAW1763 Pepti 121.5 4.0 1102 21 AAG43349 AAAB199 Canin 120.5 3.9 2665 22 AAW6550 Pepti 120.5 3.9 2665 22 AAW6550 PEDI 120.5 3.9 2665 22 AAW65050 PEDI 120.5 3.9 2665 22 AAW6650 PEDI 120.5 3.0 2665 22 AAW6650 PEDI 120.5 3.0 2665 22 AAW66	137 4.5 560 21 AAB19182 Lipid Mouse 131.5 4.3 291 20 AAV141200 Mouse 131.5 4.3 291 20 AAV141200 Mouse 131.5 4.3 291 20 AAV14120 Human 130.5 4.3 411 22 AAG4331 Human 130.5 4.3 411 22 AAG64314 Human 130.5 4.3 411 22 AAG64314 Human 130.5 4.3 411 22 AAG64314 Human 120.5 4.3 411 22 AAG64314 Human 120.5 4.3 867 19 AAV898564 Human 127.5 4.2 281 20 AAV48439 Human 127.5 4.2 281 20 AAV48439 Human 120.5 4.0 120.2 AAV17698 Mouse 120.5 4.0 120.2 AAV17693 AAV1763
11.5 4.6 7818 2.0 AANUG9555 HINBAN HINBAN AANUG951	11.5 4.7 560 21 AANUG9355 HINBAN 13.5 4.5 560 21 AANUG9355 HOLDYD 13.5 4.3 285 20 AANT4200 Mouse 13.5 4.3 285 20 AANT4200 Mouse 13.5 4.3 281 20 AANT4198 Human 130.5 4.3 411 22 AAG89915 Human 127.5 4.2 281 20 AAV98879 Human 127.5 4.2 281 20 AAV48439 Human 127.5 4.2 281 20 AAV48439 Human 126 4.1 263 19 AAV14980 Mouse 122.5 4.0 120.2 AAMT7663 Human 122.5 4.0 120.2 AAMT7663 AARD5136 AARD51
145 4.7 2329 18 AAVZ5038 PARCAZ 143 4.7 3418 20 AAV04358 Human 141.5 4.6 560 21 AAR10182 POLYP 131.5 4.3 285 20 AAV14200 POLYP 131.5 4.3 285 20 AAV14198 PUNDS 131.5 4.3 754 19 AAV33811 PARM33 130.5 4.3 411 22 AAS6484 PUMBAN 130.5 4.3 411 22 AAS6484 PUMBAN 130.6 4.3 411 22 AAS6484 PUMBAN 127.5 4.2 281 20 AAV3829 PUMBAN 126 4.1 263 19 AAV34849 PUMBAN 126 4.1 263 19 AAV34849 PUMBAN 127.5 4.0 1202 22 AAV37639 POPCHI 121.5 4.0 1024 21 AAG42350 AAR3N 121.5 4.0 1105 21 AAG42349 AAR3N 120.5 3.9 2665 22 AAV36989 Canin 120.5 3.9 2665 22 AAV36950 POPCHI 120.5 3.9 2665 22 AAV36950 POPCHI 120.5 3.9 2665 22 AAV36950 POPCHI	145 4.7 2329 18 AAVZ5038 PARCAZ 143 4.7 3418 20 AAV04355 POLTUP 141.5 4.6 582 21 AAVZ14200 POLTUP 131.5 4.3 285 20 AAV14200 POLTUP 131.5 4.3 281 20 AAV14198 PHUMBA 131.5 4.3 281 20 AAV14198 PHUMBA 130.5 4.3 411 22 AAB99915 PHUMBA 130.6 4.3 411 22 AAB9864 PHUMBA 130.7 4.3 411 22 AAB9864 PHUMBA 130.8 4.3 411 22 AAB98664 PHUMBA 130.9 4.3 867 19 AAW98879 PHUMBA 127.5 4.2 281 20 AAV73829 PHUMBA 127.5 4.2 281 20 AAV74980 PHUMBA 127.5 4.0 1202 22 AAW17663 PEPTIF 121.5 4.0 1024 21 AAG42348 AAVABB 121.5 4.0 1024 21 AAG42348 AAVABB 120.5 3.9 2343 20 AAW80989 AAVABB 120.5 4.0 1024 21 AAG42348 AAVABB 120.5 4.0 1189 21 AAG42348 AAVABB 120.5 3.9 2343 20 AAW80989 Canin
145 4.7 3418 20 AA174359 Human 144 4.7 3418 21 AA777819 Human 141.5 4.7 3418 21 AA777819 Human 141.5 4.7 3418 21 AA777819 Human 141.5 4.6 782 11 AAR06991 Human 141.5 4.6 782 11 AA806991 Human 141.5 4.6 782 11 AA806991 Human 141.5 4.6 782 11 AA806991 Human 141.5 4.3 285 21 AA819182 Lipid Human 130.5 4.3 411 22 AA64314 Human 120.5 4.0 20 AAW00005 Human 127.5 4.2 281 20 AA773829 Human 127.5 4.2 281 20 AA744980 Human 122.5 4.0 1202 22 AAM1763 Human 122.5 4.0 1202 22 AAM1763 Human 122.5 4.0 1024 21 AA642349 AA741199 AA7411121.5 4.0 1189 21 AA642349 AA741133 AA741133 AA741133 AA741133 AA741133 AA741133 AA741133 AA741133 AA741133 AA741134 AA741133 AA741134 AA74133 AA74	145 4.7 3418 20 AAV04359 Human 145 4.7 3418 20 AAV04359 Human 144 4.7 3418 21 AAV27819 BRCA2 144 4.7 3418 21 AAV25038 Human 144 4.7 3418 20 AAV36038 Human 13.5 4.6 782 11 AAR06091 POJPP 13.5 4.5 285 20 AAV14200 Mouse 2 13.5 4.3 281 20 AAV14109 Human 13.5 4.3 281 20 AAV14109 Human 5 130.5 4.3 411 22 AA664314 Human 5 130.5 4.3 411 22 AA664314 Human 6 130.5 4.3 411 22 AA664314 Human 120 AAV398005 Human 120.5 4.0 20 AAV49879 Human 120.5 4.0 20 AAV4980 Mouse 2 AAV47480 Mouse 3 AAV47480 Mous
145 4.7 3418 20 AAY04356 Human BRCA2 445 447 3418 20 AAY04357 Human BRCA2 445 447 3418 20 AAY04354 Human BRCA2 445 447 3418 21 AAY07319 BRCA2 BRCA2 BRCA2 BRCA2 DAY07318 21 2329 18 AAW25038 BRCA2 BRCA2 DAY07319 BACA2 BRCA2 BRCA2 DAY07319 D	145 4.7 3418 20 AAV04355 Hinman 145 4.7 3418 20 AAV04356 Hinman 145 4.7 3418 20 AAV04356 Hinman 145 4.7 3418 20 AAV04354 Hinman 145 4.7 3418 21 AAV02363 Hinman 144 4.7 2329 18 AAW25038 Hinman 143 5 4.5 560 21 AAR06991 Lipid Polypous 13.5 4.3 291 20 AAV14190 Hinman 131.5 4.3 291 20 AAV14190 Hinman 130.5 4.3 411 22 AA66434 Hinman 120.5 4.0 281 20 AAV74880 Hinman 120.5 4.0 281 20 AAV74840 Hinman 120.5 4.0 120.2 22 AAW06223 Hinman 120.5 4.0 120.2 22 AAW06234 Hinm

```
This sequence is the human MED1 endonuclease of the invention.

MED1 (for methyl-CpG binding endonuclease-1) is used to screen for specific modulators (potential therapeutic agents particularly minetics of MED1) and to study interactions involved in maintaining DNA fidelity.

Cor DNA manipulation and to raise antibodies. Susceptibility or predisposition to cancer (particularly colorectal or endometrial, especially hereditary non-polyposis colorectal or endometrial, especially hereditary non-polyposis colorectal cancer), or its prognosis, where caused by alterations in the MED1-encoding gene, are identified by sequence comparison, amplification, detecting altered polypeptide, and restriction fragment mapping, hybridisation (particularly to probes specific for a mutant allele). These same methods can also be used to diagnose fragile X syndrome and other diseases (e.g. myotonic dystrophy, Huntington's disease, spinocerebellar ataxia and Kennedy's disease)

Suscitated with triplet repeat expansion. The DNA, or its fragments, are used as probes and primers in the above diagnostic methods, also to also and primers in the above diagnostic methods, also to cancer calls.
fragile X syndrome, and therapeutically
                                                       Claim 8; Fig 3; 109pp; English.
```

Human delta228-UV damage endonuclease.

(first entry)

27-MAR-2000

AAY44504;

AAY44504 standard; Protein; 565 AA.

580 AA; Sequence

```
0
                                                                                                                                                                                                                                                                                                                              241 ipikktkkgcrkscsgfvqsdskresvcnkadaesepvaqksqldrtvcisdagacgetl 300
                                                                                                                                                                                                                                                                                                                                                                                             121 KSSLANYLHKNGETSLKPEDFDFTVLSKRGIKSRYKDCSMAALTSHLQNQSNNSNWNLRT 180
                                                                                                                                                                                                       181 RSKCKKDVFMPPSSSSELQESRGLSNFTSTHLLLKEDEGVDDVNFRKVRKPKGKVTILKG 240
                                                                                                                                                                                                                                                                   241 IPIKKTKKGCRKSCSGFVQSDSKRESVCNKADAESEPVAQKSQLDRTVCISDAGACGETL 300
                                                                                                                                                                                                                                                                                                                                                                            SVTSEENSLVKKKERSLSSGSNFCSEQKTSGIINKFCSAKDSEHNEKYEDTFLESEEIGT 360
                                                                                                                                                                                                                                                                                                                                                                                                                                        361 KVEVVERKEHLHTDILKRGSEMDNNCSPTRKDFTGEKIFQEDTIPRTQIERRKTSLYFSS 420
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   KYNKEALSPPRRKAFKKWTPPRSPFNLVQETLFHDPWKLLIATIFLNRTSGKMAIPVLWK 480
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              481 FLEKYPSAEVARTADWRDVSELLKPLGLYDLRAKTIVKFSDEYLTKOWKYPIELHGIGKY 540
                                                                                                                          61 CNPLLQEPIASAQFGATAGTECRKSVPCGWERVVKQRLFGKTAGRFDVYFISPQGLKFRS 120
                               0; Gaps
                                                              1 MGTTGLESLSLGDRGAAPTVTSSERLVPDPPNDLRKEDVAMELERVGEDEEQMMIKRSSE 60
                                                                                            1 mgttgles1s1gdrgaaptvtsser1vpdppnd1rkedvamelervgedeeqmmikrsse 60
                                                                             100.0%; Score 3055; DB 20; Length 580; 100.0%; Pred. No. 7.6e-284;
                                0; Indels
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           541 GNDSYRIFCVNEWKQVHPEDHKLNKYHDWLWENHEKLSLS 580
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            0; Mismatches
           watches 580; Conservative
   Query Match
                                                                                                                                                                                                                                                                                                                                                                            301
                                                                                                                                                                                                                                                                                                                                                                                                          301
                                                                δ
                                                                                            Dp
                                                                                                                          δλ
                                                                                                                                                       g
                                                                                                                                                                                                                   g
                                                                                                                                                                                                                                                  δy
                                                                                                                                                                                                                                                                                g
                                                                                                                                                                                                                                                                                                             Qγ
                                                                                                                                                                                                                                                                                                                                         Db
                                                                                                                                                                                                                                                                                                                                                                            ò
                                                                                                                                                                                                                                                                                                                                                                                                       g
                                                                                                                                                                                                                                                                                                                                                                                                                                    Óγ
                                                                                                                                                                                                                                                                                                                                                                                                                                                                    q
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 δ
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             QQ
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              δy
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          qq
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          δ
```

AAY44504 RESULT

```
The present sequence is human delta228-UV damage endonuclease. Delta228-UVDE is a 288 amino acid deletion of the N-terminal of the uvel+ gene product. This is expressed in frame with a GST leader sequence
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      structure resulting from UV radiation damage, a photoproduct, an abasic site, misnatched nucleotide pairing, aplathum diaduct, an intercalated molecule or alkylation of a nucleotide. Uvelp can be used in compositions for internal or topical application and as a therapeutic agent for skin
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   to generate a fusion protein. This provides stable endonuclease fragments for cleaving a double-stranded DNA molecule that has a distorted
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       A new truncated ultraviolet damage endonuclease for treatment of skin
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       61 cnpllqepiasaqfgatagtecrksvpcgwervvkqr]fgktagrfdvyfispqglkfrs 120
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               181 RSKCKKDVFMPPSSSSELQESRGLSNFTSTHLLLKEDEGVDDVNFRKVRKPKGKVTILKG 240
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       61 CNPLLQEPIASAQFGATAGTECRKSVPCGWERVVKQRLFGKTAGRFDVYFISPQGLKFRS 120
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           121 KSSLANYLHKNGETSLKPEDFDFTVLSKRGIKSRYKDCSMAALTSHLONQSNNSNWNLRT 180
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           241 TPTKKTKKGCRKSCSGFVQSDSKRESVCNKADAESEPVAQKSQLDRTVCISDAGACGETL 300
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      96.8%; Score 2958.5; DB 21; Length 565;
97.4%; Pred. No. 1.3e-274;
Live 0; Mismatches 0; Indels 15; Gaps
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                1 MGTTGLESLSLGDRGAAPTVTSSERLVPDPPNDLRKEDVAMELERVGEDEEQMMIKRSSE 60
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   1 mgttgles.lslgdrgaaptvtsserlvpdppndlrkedvamelervgedeegmmikrsse 60
                                                                                                                                             Delta228-UVDE; ultraviolet damage endonuclease; GST signal peptide;
                                                                                                                                                             glutathione-S-transferase signal peptide, uvel+ gene product; UV irradiation; DNA damage; UV radiation damage; photoproduct; abasic site; aplatinum diaduct; mismatched nucleotide pairing;
                                                                                                                                                                                                                   nucleotide alkylation; skin cancer.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            claim 16; Page 60; 133pp; English.
                                                                                                                                                                                                                                                                                                                                                                                                      98US-0088521.
                                                                                                                                                                                                                                                                                                                                                                     99WO-US12910
                                                                                                                                                                                                                                                                                                                                                                                                                        99US-0134752
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         Best Local Similarity 97.4 Matches 565; Conservative
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                Kaur B,
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   WPI; 2000-116417/10.
                                                                                                                                                                                                                                                                                                                                                                                                                                                             (UYEM-) UNIV EMORY.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   565 AA;
                                                                                                                                                                                                                                                            Homo sapiens.
                                                                                                                                                                                                                                                                                               WO9963828-A1.
                                                                                                                                                                                                                                                                                                                                                                   08-JUN-1999;
                                                                                                                                                                                                                                                                                                                                                                                                        08-JUN-1998;
                                                                                                                                                                                                                                                                                                                                                                                                                        18-MAY-1999;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                Doetsch PW,
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     Sequence
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          Query Match
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    cancers.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           cancers
δò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       qq
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           QΥ
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 ô
```

ij

fragments encoded by the human ovarian tumor cDNA library derived EST fragments represented in AA277450-277572.

439 AA;

Sequence

22 × 88

```
301 SVTSEENSLVKKKERSLSSGSNFCSEQKTSGIINKFCSAKDSEHNEKYEDTFLESEEIGT 360
                                                      420
                                                                                  405
                                                                                                           KYNKEALSPPRRKAFKKWTPPRSPFNLVQETLFHDPWKLLIATIFLNRTSGKMAIPVLWK 480
                                                                                                                                       465
                                                                                                                                                                              flekypsaevartadwrdvsellkplglydlraktivkfsdeyltkqwkypielhgigky 525
                                                                     FLEKYPSAEVARTADWRDVSELLKPLGLYDLRAKTIVKFSDEYLTKQWKYPIELHGIGKY
                                                      KVEVVERKEHLHTDILKRGSEMDNNCSPTRKDFTGEKLFQEDTIPRTQIERRKTSLYFSS
                                                                                                                                                                                                                       541 GNDSYRIFCVNEWKQVHPEDHKLNKYHDWLWENHEKLSLS 580
                                                                                                                                                                                                                                      526 gndsyrifcvnewkqvhpedhklnkyhdwlwenhekls1s (565)
                                                                                                                                                                                            466
                          289
                                                    361
                                                                               346
                                                                                                           421
                                                                                                                                                                481
                                                                                                                                    셤
                                                                                                                                                                                          g
                                                                                                                                                                                                                       ò
                                                                                                                                                                                                                                                  9
                                                    δ
                                                                             8
                                                                                                           ò
                                                                                                                                                                 ŏ
```

Expressed sequence tag; EST; human; ovarian tumor; anticancer; Human ovarian tumor EST fragment encoded protein 44. AAY76548 standard; Protein; 439 AA. gene therapy; treatment. DE19817557-A1 Homo sapiens 10-APR-2000 AAY76548 AAY76548 NX K

(META-) METAGEN GES GENOMFORSCHUNG MBH. 98DE-1017557 98DE-1017557 09-APR-1998; 09-APR-1998; 21-OCT-1999

Ξ Dah. Rosenthal A, Specht T, Hinzmann B, Schmitt A, Pilarsky C, WPI; 1999-591920/51.

N-PSDB; AAZ77469

New nucleic acid sequences expressed in ovarian, and some other, cancer tissues, and derived polypeptides, for treatment of ovarian cancer and identification of therapeutic agents

Claim 25; Page 261; 310pp; German

have anticancer activity and are highly expressed in ovarian tumor tissue (and some also in testis and breast cancer tissue). The products of the invention can be used for gene therapy. (A) are used (1) for recombinant expression of polypeptides (B) and (ii) to isolate complete genes. (B) are used (i) to identify agents suitable for treatment of ovarian cancer; (ii) directly for treating this form of cancer (including expression from gene therapy vectors) and (iii) for generation of specific antibodies. (A) are identified by assembling ESTs (expressed sequence tags) from a particular tissue type before comparison of expression patterns. This allows a significantly longer fragment of the expression patterns. This allows a significantly longer fragment of the caper to be revealed, so should reduce the number of failures associated with the fact that ESTs from different libraries may represent different parts of the same unknown gene, distorting the estimated frequency of occurrence in a particular tissue. AAY76505-Y76638 represent protein This invention describes novel nucleic acid (cDNA) sequences (A) which

97US-0048877. 97US-0048877. 97US-0048881.

06-JUN-1997; 18-DEC-1997;

98WO-US11422

04-JUN-1998;

10-DEC-1998

Homo sapiens WO9854963-A2

```
3,
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     Human; secreted protein; fusion protein; gene therapy; protein therapy; diagnosis; tissue; cancer; tumour; neurodegenerative disorder; leukaemia; developmental abnormality; foetal deficiency; blood; allergy; renal; immune system; asthma; lymphocytic disease; brain; hepatic; lymphoma; inflammation; ischaemic shock; Alzheimer's disease; restenosis; AIDS; osquitive disorder; schizophrenia; prostate; obssity; osteoclast; thymus; osteoplast; arthritis; testis; lung; thyroidtis; thyroid; digestion; endocrine; metabolism; regulation; malabsorption; gastritis; neoplasm.
                                                                95 KQRLFGKTAGRFD-VYFISPQGLKFRSKSSLANYLHKNGETSLKPEDFDFTVLSKRGIKS 153
                                                                                                                                 RYKDCSMAALTSHLQNQSNNSNWNLRTRSKCKKDVFMPPSSSSELQESRGLSNFTSTHLL 213
                                                                                                                                                                 rykdcsmaaltshlqnqsnnsnwnlrtrskckkdvfmppsssselqesrglsnftsthll 118
                                                                                                                                                                                                  214 LKEDEGVDDVNFRKVRKPKGKVT1LKGIPIKKTKKGCRKSCSGFVQSDSKRESVCNKADA 273
                                                                                                                                                                                                                   ESEPVAQKSQLDRTVCISDAGACGETLSVTSEENSLVKKKERSLSSGSNFCSEQKTSGII 333
                                                                                                                                                                                                                                                                                     TGEKIFQEDT1PRTQIERRKTSLYFSSKYNKEALSPPRRKAFKKWTPPRSPFNLVQETLF 453
                                                                                                                                                                                                                                                                                                                                                                                                                                      t----edtiprtgierrktslyfsskynkealspprrkafkkwtpprspfnlygetlf 352
                                                                                                                                                                                                                                                                                                                                                                                                                                                                     HDPWKLLIATIFLNRTSGKMAIPVLWKFLEKYPSAEVARTADWRDVSELLKPLGLYDLRA 513
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     9; Gaps
                                                                                   NKFCSAKDSEHNEKYEDTFLESEEIGTKVEVVERKEHLHTDILKRGSEMDNNCSPTRKDF
   Length 439;
Score 2194.5; DB 20; Lengt
Pred. No. 1.6e-201;
3; Mismatches 6; Indels
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           Secreted protein encoded by gene 168 clone HCFNF11.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      AAW88701 standard; Protein; 257 AA
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      514 KTIVKFSDEYLTKQWKYPIELHGIG 538
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     413 ktivkfsdeyltkqwkypielhgig 437
 71.88;
96.08;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      01-MAR-1999 (first entry)
                               Matches 427; Conservative
 Query Match
Best Local Similarity
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         AAW88701;
                                                                                                                                                                                                                                                                                                                                                                    239
                                                                                                                                                                                                                                                                                                                                                                                                     394
                                                                                                                                                                                                                                                                                                                                                                                                                                                                     454
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     353
                                                                                                                                                                 59
                                                                                                                                                                                                                                                                   274
                                                                                                                                                                                                                                                                                                                                     334
                                                                                                                                                                                                                                                                                                                                                                                                                                     299
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        AAW88701
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       RESULT
                                                                 δŏ
                                                                                                 q
                                                                                                                                 Qγ
                                                                                                                                                               qq
                                                                                                                                                                                                  ŏ
                                                                                                                                                                                                                                 q
                                                                                                                                                                                                                                                                   ô
                                                                                                                                                                                                                                                                                                 qq
                                                                                                                                                                                                                                                                                                                                    ò
                                                                                                                                                                                                                                                                                                                                                                   qq
                                                                                                                                                                                                                                                                                                                                                                                                    δy
                                                                                                                                                                                                                                                                                                                                                                                                                                     a
                                                                                                                                                                                                                                                                                                                                                                                                                                                                     à
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     a
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      QY
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     qq
```

```
97US-0057650.
97US-0057661.
97US-0057667.
97US-0057761.
                                        97US-0048899.
97US-0048915.
97US-0048949.
97US-0048964.
                                                                                                                                                                                                                                                                 97US-0057764.
97US-0057770.
97US-0057775.
                                                                                                                                                                                                                                                                                                              97US-0057778.
97US-0048875.
97US-0048878.
                                                                                                                                                                                                                                                                                                                                                                                                                     97US-0048900.
97US-0048916.
97US-0048962.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                 97US-0048970.
97US-0048974.
97US-0049373.
                                                                                                                                                                                            97US-0057647.
                                                                                                                                                                                                                                                                                                                                                           97US-0048882.
97US-0048885.
                                                                                                                                                                                                                                                                                                                                                                                         97US-0048894.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      97US-0057648.
97US-0057651.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   97US-0057668.
97US-0057762.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                97US-0057765.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           97US-0048880.
97US-0048883.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     97US-0048901.
97US-0048917.
                                                                                                                                                                                                                                                                                                                                                                                                      97US-0048897.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                97US-0057584.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             97US-0057642.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          97US-0057645.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                97US-0057776.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                97US-0057777.
                                                                                                                                                 97US-0057628.
97US-0057635.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              97US-0048876.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          97US-0048892.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  97US-0048971.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           97US-0057646.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        97US-0057649.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        97US-0057654.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       97us-0057666.
                                                                                                                                 97US-0049375
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  97US-0049019
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               97US-0049374
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           97US-0057627
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              97us-0057771
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           97US-0057643
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                97US-0057769
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             97US-0057
                                                                                                                                                                            97US-0057
                                        06-JUN-1997

06-JUN-1997

06-JUN-1997

06-JUN-1997

05-SEP-1997

05-SEP-1997

05-SEP-1997

05-SEP-1997

05-SEP-1997

05-SEP-1997

05-SEP-1997

05-SEP-1997

06-JUN-1997

07-JUN-1997

08-JUN-1997

08-JUN-1997

08-JUN-1997

08-JUN-1997

08-JUN-1997

08-JUN-1997

08-JUN-1997

08-JUN-1997

08-JU
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     05-SEP-1997;
05-SEP-1997;
05-SEP-1997;
05-SEP-1997;
05-SEP-1997;
05-SEP-1997;
```

(HUMA-) HUMAN GENOME SCI INC

```
The invention relates to nucleic acid sequences (AAV84411 to AAV84633)
encoding human secreted proteins (AAW88534 to AAW8876). The secreted
protein gene sequences are deposited with the Arcu under deposit numbers
crossing the sequences are deposited with the Arcu under deposit numbers
arc 97979, 97974, 97975, 97976, 209082, 209085, 209085, 209085, 209081, 2090011, 2090011, 2090081, 2090081, 2090081, 2090085, 2090085, 2090085, 209511. Host
cells comprising recombinant vectors containing the nucleic acid
sequences are used for the recombinant production of the secreted
proteins. The polynucleotide and amino acid sequences are useful for are
useful for preventing, treating or ameliorating medical conditions e.g.
by protein or gene therapy. Pathological conditions can be also
diagnosed by determining the amount of the new polypeptides in a sample
or by determining the presence of mutations in the new polynucleotides.
Specific uses are described for each of the polynucleotides, based on
which tissues they are most highly expressed in, and include developing
disorders, developmental abnormalities and foetal deficiencies, blood
disorders, tennours, leukentas, diseases of the immune system, autoingmune
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        diseases, heparic and renal disease, lymphomas, inflammation, allergies, ischemic shock, Alzheimer's and cognitive disorders, schizophrenia, restenosis, prostate diseases, obesity, disorders involving osteoclasts such as osteoporosis, arthritis or malignancies, diseases of testes, lung or thymus, digestive/endocrine disorders, infections and AIDS. The polypeptides are also useful for identifying their binding partners. The present sequence represents human secreted protein (see descriptor line for gene number and clone identification).
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           160 MAALTSHLQNQSNNSNWNLRTRSKCKKDVFMPPSSSSELQESRGLSNFTSTHLLLKEDEG 219
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                220 VDDVNFRKVRKPKGKVTILKGIPIKKTKKGCRKSCSGFVQSDSKRESVCNKADAESEPVA 279
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    280 QKSQLDRTVC1SDAGACGETLSVTSEENSLVKKKERSLSSGSNFCSEQKTSGIINKFCSA 339
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      340 KDSEHNEKYEDTFLESEEIGTKVEVVERKEHLHTDILKRGSEMDNNCSPTRKDFTGEKIF 399
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          6; Gaps
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               1 maaltshlqnqsnnsnwnlrtrskckkdvfmppsssselqesrglsnftsthlilkedeg 60
                                                                                                                                                                                                                                  New isolated human genes and the secreted polypeptides they encode useful for diagnosis and treatment of e.g. cancers, neurological disorders, immune diseases, inflammation or blood disorders
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   40.3%; Score 1230; DB 20; Length 257; 93.1%; Pred. No. 2.1e-109; ive 3; Mismatches 9; Indels 6;
Brewer LA, Carter KC, Dillon PJ, Ebner R, Endress GA, Fran P, Feng P, Ferrie AM, Fischer CL, Florence C; Florence K, Greene JM, Hu J, Kyaw H, Lafleur DM; Li Y, Moore PA, Ni J, Olsen HS, Rosen CA, Ruben SM; Shi Y, Soppet DR, Wei Y, Young P, Yu G, Zeng Z;
                                                                                                                                                                                                                                                                                                                                         Claim 11; Page 564-565; 772pp; English.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       AAG02051 standard; Protein; 219 AA.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           400 QEDTIPRTQIERRKTSLYF 418
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 236 -edtiprntdrkkenkpvf 253
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       Matches 241; Conservative
                                                                                                                                                           WPI; 1999-059865/05.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            Best Local Similarity
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             257 AA;
                                                                                                                                                                                      N-PSDB; AAV84578.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  Sednence
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            Query Match
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    AAG02051
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           RESULL
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         OI XX
     δ
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                ŏλ
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 q
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                a
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    ΟŊ
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       qq
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           Öλ
```

ï

```
18-JUN-1999;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    -APR-]
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             30-APR-1
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    04-MAY-1
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       05-MAY-
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  3-JUN-
                                                                                                                                                                                                                                                                                                                                                                                                                                                           06-APR-
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       -APR-
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    23-APR-
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  30-APR-
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      -NDC-01
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        21-APR-
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                33-APR-
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         28-APR-
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         28-MAY-
'n
                                                                                                                    Human; 5' EST; expressed sequence tag; secreted protein; cDNA isolation; gene therapy; chromosome mapping.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             The present sequence is a polypeptide encoded by one of a large number of 5' ESTs derived from mRNAs encoding secreted proteins. The 5' ESTs different tissues. EST sequences usually correspond mainly to the 3' untranslated region (UTR) of the mRNA because they are often obtained from oligo-dT primed cDNA libraries. Such ESTs are not well suited for isolating cDNA sequences derived from the 5' ends of mRNAs and even in those cases where longer cDNA sequences have been obtained, the full 5' UTR is rarely included. 5' ESTs are derived from mRNAs with intact 5' ends and can therefore be used to obtain full length cDNAs and genomic DNAs. 5' ESTs are also used in diagnostic, forensic, gene therapy and chromosome mapping procedures. They are used to obtain upstream regulatory sequences and to design expression and secretion vectors.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             New nucleic acid that is a 5' expressed sequence tag (5' EST) for obtaining cDNAs and genomic DNAs that correspond to 5'ESTs and for diagnostic, forensic, gene therapy and chromosome mapping procedures
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           ECRKSV------PCGWERVVKQRLFGKTAGRFDVYFISPQGLKFRSKSSLANY 127
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  47; Indels 30; Gaps
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           EDVAMELERVGEDEEQMMIKRSSECNPLLQEPIASAQFG------ATAGT 80
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  34 kdxplkfxxvkkdkkeekegkhepvqpsahhsaepaeagkaetsxgsgsapavpeasasp 93
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     6.4%; Score 194; DB 21; Length 219; 33.3%; Pred. No. 3.7e-10; tive 21; Mismatches 47; Indels 36
                                                                                                                                                                                                                                                                                                                                                                                                                                       Dumas Milne Edwards J, Duclert A, Giordano J;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             Claim 13; SEQ ID 6132; 71pp + CD-ROM; English
                                                                               Human secreted protein, SEQ ID NO: 6132.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 128 LHKNGETSLKPFPFPFTVLSKRGIKSR 154
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        AAG16859 standard; Protein; 326 AA.
                                                                                                                                                                                                                                                                                                          21-FEB-2000; 2000EP-0200610.
                                                                                                                                                                                                                                                                                                                                                     99US-0122487
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           Best Local Similarity 33.3% Matches 49; Conservative
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          17-OCT-2000 (first entry)
                                    06-OCT-2000 (first entry)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                              WPI; 2000-500381/45.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               219 AA;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    N-PSDB; AAC02057
                                                                                                                                                                                                                                                                                                                                                                                              (GEST ) GENSET
                                                                                                                                                                                                                                                                                                                                                     26-FEB-1999;
                                                                                                                                                                                     Homo sapiens
                                                                                                                                                                                                                            EP1033401-A2
                                                                                                                                                                                                                                                                     06-SEP-2000
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               Sequence
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 AAG16859;
AAG02051;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       Query Match
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         37
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    AAG16859
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                qq
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           δ
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            G
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                Y X X Y
NAME OF COLOR OF STREET OF
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           ç
```

```
Protein identification; signal transduction pathway; metabolic pathway; hybridisation assay; genetic mapping; gene expression control; promoter; termination sequence.
Arabidopsis thaliana protein fragment SEQ ID NO: 17663.
                                                                                                                                                                                                                                                                                                                                                                                                                                                 99US-0134218.
99US-0134218.
99US-0134219.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               99US-0139463.
99US-0139750.
99US-0139763.
                                                                                                                                                                                                                                                                                     29845.
                                                                                                                                                                                                                                                                                                            99US-0130449.
                                                                                                                                                                                                                                                                                                                                     99US-0130891.
                                                                                                                                                                                                                                                                                                                                                             99US-0132048.
                                                                                                                                                                                                                                                                                                                                                                                                99US-0132485.
                                                                                                                                                                                                                                                                                                                                                                                                                          99US-0132487.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          99US-0136021.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             99US-0137222
99US-0137528
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       99US-0138847
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            99US-0139456.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    99US-0139458
99US-0139459
                                                                                                                                                                                               99US-0123548.
                                                                                                                                                                                                                                                99US-0127462.
99US-0128234.
                                                                                                                                                                                                                                                                          99US-0128714
                                                                                                                                                                                                                                                                                                                                                                                       99US-0132484
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      99US-0134221
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   99US-0134370
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          99US-0134941
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        99US-0135124
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   99US-0135353
99US-0135629
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      99US-0137502
99US-0137724
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               99US-0138094
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           99US-0138540
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    99US-0139454
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              99US-0139460
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      990S-0139462
                                                                                                                                               2000EP-0301439
                                                                                                                                                                        99US-0121825
                                                                                                                                                                                     99US-0123180
                                                                                                                                                                                                                         99US-0126264
                                                                                                                                                                                                                                      99US-0126785
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                99US-0134768
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   99US-0136782
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 99US-0139452
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              99US-0139453
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         99US-0139492
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  99US-0139455
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         99US-0139461
                                                                                                                                                                                                                                                                                                  99US-01
                                                                                                                                                                                                                                                                                      99US-01
                                                                       Arabidopsis thaliana
                                                                                                                                                                                                                                                                                                                                                                                                                         06-MAY-1999;
07-MAY-1999;
11-MAY-1999;
14-MAY-1999;
14-MAY-1999;
14-MAY-1999;
18-MAY-1999;
20-MAY-1999;
21-MAY-1999;
21-MAY-1999;
21-MAY-1999;
21-MAY-1999;
                                                                                                                                               25-FEB-2000;
                                                                                                 EP1033405-A2
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        1999;
                                                                                                                                                                                                                                                                                                                                     1999;
                                                                                                                                                                                                                                                                                                                                                             1999;
                                                                                                                                                                                                                                                                                                                                                                                                  16661
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    1999;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       1999;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  14-JUN-1999;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    (8-JUN-1999;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        18-JUN-1999;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          25-MAY-1999;
27-MAY-1999;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              10-JUN-1999;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               08-JUN-1999;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            JUN-1999;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 16-JUN-1999;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            16-JUN-1999;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         7-JUN-1999;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  .8-JUN-1999;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            18-JUN-1999;
                                                                                                                        06-SEP-2000.
                                                                                                                                                                                                            23-MAR-1999;
25-MAR-1999;
                                                                                                                                                                                                                                      29-MAR-1999;
                                                                                                                                                                                                                                                  01-APR-1999;
                                                                                                                                                                                                                                                                          08-APR-1999;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    07-JUN-1999;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            18-JUN-1999
                                                                                                                                                                                                                                                                                                                                                                                                               - MAY - 1
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       04 - JUN-1
```

90S - 0139817 90S - 0139899 90S - 0140354 90S - 0140695 90S - 0140695 90S - 0141287 90S - 0141287 90S - 0141287 90S - 0142890 90S - 0142390 90S - 0142390 90S - 0142390	990S-0143542 990S-0144085 990S-0144086 990S-0144086 990S-01441331 990S-01441331 990S-01441331 990S-01441332 990S-01441332 990S-01441334 990S-01441336 990S-01441336 990S-01441336 990S-01441336 990S-0146132 990S-0145086 990S-0145086 990S-0145086 990S-0145086 990S-0145196 990S-0145196 990S-0145196 990S-0145196 990S-0145196 990S-0145918 990S-0145918 990S-0145918 990S-0145918 990S-0145918 990S-0145918 990S-0145918 990S-0145918 990S-0145918 990S-0145918 990S-0145918 990S-0145918	0.03 - 0.14 74 93 3 0.03 - 0.14 74 93 3 0.03 - 0.14 83 19 0.03 - 0.14 86 84 0.03 - 0.14 86 84 0.03 - 0.14 91 25 0.03 - 0.14 99 20 0.03 - 0.14 99 20 0.03 - 0.14 99 20 0.03 - 0.14 99 20 0.03 - 0.14 99 20 0.03 - 0.14 99 20 0.03 - 0.14 99 20 0.03 - 0.14 99 20 0.03 - 0.14 99 20 0.03 - 0.14 99 20 0.03 - 0.14 99 20 0.03 - 0.15 99 80 0.03 - 0.15 99 80 0.03 - 0.15 90 0.03 - 0.15 90 0.03 - 0.15 90
21 JUN-1999 R 23 JUN-1999 R 23 JUN-1999 R 24 JUN-1999 R 28 JUN-1999 R 29 JUN-1999 R 01 JUL-1999 R 01 JUL-1999 R 06 JUL-1999 R 06 JUL-1999 R 06 JUL-1999 R 06 JUL-1999	PR 13-JUL-1999; PR 14-JUL-1999; PR 16-JUL-1999; PR 16-JUL-1999; PR 19-JUL-1999; PR 19-JUL-1999; PR 19-JUL-1999; PR 20-JUL-1999; PR 20-JUL-1999; PR 20-JUL-1999; PR 21-JUL-1999; PR 22-JUL-1999; PR 22-JUL-1999; PR 22-JUL-1999; PR 22-JUL-1999; PR 23-JUL-1999; PR 73-JUL-1999; PR 74-JUL-1999; PR 74-JUL-1999; PR 75-JUL-1999; PR 75-JUL-1999	R 09-AUG-1999 R 10-AUG-1999 R 11-AUG-1999 R 13-AUG-1999 R 13-AUG-1999 R 13-AUG-1999 R 20-AUG-1999 R 20-AUG-1999 R 20-AUG-1999 R 25-AUG-1999 R 25-AUG-1999 R 25-AUG-1999

```
Query Match 4.9%; Score 148.5; DB 21; Length 326; Best Local Similarity 20.7%; Pred. No. 1.6e-05; Matches 62; Conservative 47; Mismatches 102; Indels 89; Gaps
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      || | : :|: || : : | :: | 88 snlvspqiadddddsvsdshierqewsefhvevrrvspyfggstvsqqskegc----- 140
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            306 ENSLVKK-----KERSLSSGSNFCSEQKTSGIINKFCSAKDSEHNEKYEDTFLESEEIGT 360
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  205 SNFTSTHLLLKEDEGVDDVNFRKVRKPKGKVTILKGIPI------KKTKKGCRKSCSGF 257
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                258 VQSDSKRESVCNKADAES-----EPVAQKSQLDRTVCISDAGACGET----LSVTSE 305
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             361 KVEVVERKEHLHTDILKRGSEMDNNCSPTRKDFTGEKIFQEDTIPRTQIERRKTSLYFSS 420
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                226 --kvvkvsryfhadgiqvnesqkeksrnvrk-----tpivspvlslsgktddvyl-- 273
990S-0151080.
990S-0151303.
990S-0151303.
990S-0151303.
990S-0153758.
990S-0153758.
990S-0154779.
990S-0155659.
990S-0155659.
990S-0156753.
990S-0156753.
990S-0156753.
                                                                                                                                                                                                                                                                                       990S-0159584.
990S-0160741.
990S-0160767.
990S-0160768.
                                                                                                                                                                                                                                                                                                                                        990S-0160814
990S-0160815
990S-0160980
990S-0160980
990S-0161404
990S-0161406
990S-0161406
990S-0161359
990S-0161359
                                                                                                                                                                                                                               99US-0159295.
99US-0159329.
99US-0159330.
                                                                                                                                                                                                                                                           99US-0159331.
99US-0159637.
99US-0159638.
                                                                                                                                                                                                                                                                                                                                                                                                                                                   99US-0161920.
99US-0161992.
99US-0161993.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 99US-0162142
                                                                  15-SEP-1999;
16-SEP-1999;
20-SEP-1999;
22-SEP-1999;
23-SEP-1999;
24-SEP-1999;
                                                                                                                   24 - SEP - 1999;
28 - SEP - 1999;
04 - OCT - 1999;
05 - OCT - 1999;
06 - OCT - 1999;
07 - OCT - 1999;
13 - OCT - 1999;
13 - OCT - 1999;
14 - OCT - 1999;
14 - OCT - 1999;
14 - OCT - 1999;
16 - OCT - 1999;
17 - OCT - 1999;
18 - OCT - 1999;
18 - OCT - 1999;
18 - OCT - 1999;
21 - OCT - 1999;
                                                                                                                                                                                                                                                                                                                                                 21-0CT-1999)
22-0CT-1999)
22-0CT-1999)
25-0CT-1999)
25-0CT-1999)
26-0CT-1999)
26-0CT-1999)
28-0CT-1999)
28-0CT-1999)
28-0CT-1999)
                                                10-SEP-1999;
13-SEP-1999;
  qq
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       \overset{\circ}{\circ}
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     QQ
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              qq
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    Op
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         \Omega
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   οy
```

```
990s-0139457.
990s-0139458.
990s-0139469.
990s-0139461.
990s-0139461.
990s-0139463.
990s-0139463.
990s-0139463.
990s-0139817.
990s-0139817.
990s-0139817.
990s-0140853.
990s-0140823.
990s-014287.
990s-014287.
990s-014287.
990s-014287.
990s-014339.
990s-0144085.
990s-0144333.
990s-0144333.
990s-0144333.
990s-0144333.
990s-0144333.
990s-0144333.
990s-0144333.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   990S - 0145089
990S - 0145192
990S - 0145185
990S - 0145218
990S - 0145276
990S - 0145918
990S - 0145918
990S - 0145918
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             990S-0147204.
990S-0147204.
990S-0147302.
990S-0147192.
990S-0147260.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          99US-0147493.
99US-0147935.
99US-0148171.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        99US-0146386.
99US-0146388.
99US-0146389.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        99US-0149368.
99US-0149175.
99US-0149426.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              99US-0145088.
99US-0145085.
99US-0145087.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 99US-0147303.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             99US-0148565
99US-0148684
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   99US-0148319
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   990S-0148341
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            04 - AUG - 1999;
04 - AUG - 1999;
05 - AUG - 1999;
05 - AUG - 1999;
06 - AUG - 1999;
09 - AUG - 1999;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     10-AUG-1999;
11-AUG-1999;
12-AUG-1999;
13-AUG-1999;
13-AUG-1999;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        16-AUG-1999;
17-AUG-1999;
18-AUG-1999;
                                                                                                                                                                                      28-JUN-1999;
29-JUN-1999;
30-JUN-1999;
01-JUL-1999;
01-JUL-1999;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     26-JUL-1999;
27-JUL-1999;
27-JUL-1999;
27-JUL-1999;
28-JUL-1999;
18-JUN-1999;
18-JUN-1999;
18-JUN-1999;
18-JUN-1999;
18-JUN-1999;
18-JUN-1999;
18-JUN-1999;
18-JUN-1999;
23-JUN-1999;
23-JUN-1999;
23-JUN-1999;
23-JUN-1999;
                                                                                                                                                                           24-JUN-1999;
28-JUN-1999;
29-JUN-1999;
                                                                                                                                                                                                                                                                                    08-JUL-1999;
09-JUL-1999;
                                                                                                                                                                                                                                                                                                                                                                 16-JUL-1999;
16-JUL-1999;
19-JUL-1999;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    22-JUL-1999;
22-JUL-1999;
23-JUL-1999;
                                                                                                                                                                                                                                                                                                                           13-JUL-1999;
                                                                                                                                                                                                                                                                                                                                                      15-JUL-1999;
16-JUL-1999;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     21-JUL-1999;
21-JUL-1999;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           22-JUL-1999;
22-JUL-1999;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         02-AUG-1999;
02-AUG-1999;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    02-AUG-1999;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                03-AUG-1999;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            09-AUG-1999;
09-AUG-1999;
                                                                                                                                                                                                                                                          02-JUL-1999;
                                                                                                                                                                                                                                                                                                                12-JUL-1999;
                                                                                                                                                                                                                                                                                                                                                                                                                                                    19-JUL-1999;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                19-JUL-1999;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          20-JUL-1999;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         20-JUL-1999;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 21-JUL-1999;
                                                                                                                                                                                                                                                                       06-JUL-1999;
    Protein identification; signal transduction pathway; metabolic pathway; hybridisation assay; genetic mapping; gene expression control; promoter; termination sequence.
                                                                                                                                       Arabidopsis thaliana protein fragment SEQ ID NO: 17662.
                                                           AAG16858 standard; Protein; 327 AA
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               990S-0134768
990S-0135134
990S-0135134
990S-0135629-
990S-0136321
990S-0136321
990S-0137222
990S-0137222
990S-0137722
990S-0137722
990S-01377528
990S-01377528
990S-01377528
990S-013913
990S-0139452
990S-0139452
                                                                                                                                                                                                                                                                                                                             99US-0121825.
99US-0123180.
99US-0123748.
99US-0125788.
99US-0126785.
99US-0126785.
99US-0128234.
99US-01288714.
99US-0130077.
99US-0130891.
99US-0130891.
99US-0132484.
99US-0132487.
99US-0132487.
99US-0132487.
99US-0132487.
99US-0132487.
99US-0132487.
99US-0132487.
99US-0132487.
                                                                                                                                                                                                                                                                                                      25-FEB-2000; 2000EP-0301439
                                                                                                              17-OCT-2000 (first entry)
                                                                                                                                                                                                                        Arabidopsis thaliana
                                                                                                                                                                                                                                                                                                                                         05-MAR-1999;

09-MAR-1999;

23-MAR-1999;

25-MAR-1999;

01-APR-1999;

06-APR-1999;

06-APR-1999;

19-APR-1999;

23-APR-1999;

23-APR-1999;

23-APR-1999;

23-APR-1999;

06-MAY-1999;

06-MAY-1999;

06-MAY-1999;

11-MAY-1999;

11-MAY-1999;

11-MAY-1999;

14-MAY-1999;

14-MAY-1999;

19-MAY-1999;

27-MAY-1999;

21-MAY-1999;

21-MAY-1999;

21-MAY-1999;

22-MAY-1999;

22-MAY-1999;

23-MAY-1999;

24-MAY-1999;

27-MAY-1999;

27-MAY-1999;

28-MAY-1999;

27-MAY-1999;

27-MAY-1999;

28-MAY-1999;

28-MAY-1999;

21-MAY-1999;

21-MAY-1999;

21-MAY-1999;

21-MAY-1999;

21-MAY-1999;

21-MAY-1999;

21-MAY-1999;

21-MAY-1999;

21-MAY-1999;

21-MAY-1999;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             08 JUN 1999)
10 JUN 1999)
10 JUN 1999)
14 JUN 1999)
16 JUN 1999)
17 JUN 1999)
17 JUN 1999)
18 JUN 1999)
18 JUN 1999)
                                                                                                                                                                                                                                                  EP1033405-A2.
                                                                                                                                                                                                                                                                           06-SEP-2000
                                                                                                                                                                                                                                                                                                                                 25-FEB-1999
                                                                                    AAG16858;
                                RESULT
AAG16858

AAG16858

XXX
AAC66858

XXX
AAC66858

A
```

```
AAG16857 standard; Protein; 352 AA
                                                                                                                                                                                                                                                                                                                                                                                        990S-0132407.
990S-0132484.
990S-0132485.
990S-0132486.
                                                                                                                                                                                                                                                                                                                                                                                                                              990S-0132863.
99US-0134256.
99US-0134218.
99US-0134219.
99US-0134221.
                                                                                                                                                                                                                                                                                                                            99US-0128714.
99US-0129845.
99US-0130077.
                                                                                                                                                                                                                                                               99US-0121825.
99US-0123180.
99US-0123548.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  990S-0134941.
990S-0135124.
990S-0135353.
990S-0135629.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               990S-0137528.
990S-0137502.
990S-0137724.
990S-0138094.
                                                                                                                                                                                                                                                                                                                                                                   99US-0130891.
99US-0131449.
                                                                                                                                                                                                                                                                                      99US-0125788
                                                                                                                                                                                                                                                                                                      99US-0126785.
99US-0127462.
                                                                                                                                                                                                                                                                                                                                                   99US-0130449.
                                                                                                                                                                                                                                                                                                                                                           99US-0130510.
                                                                                                                                                                                                                                                                                                                                                                                   99US-0132048.
                                                                                                                                                                                                                                                                                                                     99US-0128234
                                                                                                                                                                                                                                                25-FEB-2000; 2000EP-0301439
                                                                                                                                                                                                                                                                                                                                                                                                                                                                             99US-0134768
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   99US-0136021.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         99us-0136392
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          99US-0137222
                                                                                                                                      (first entry)
                                                                                                                                                                                                   Arabidopsis thaliana.
                                                                                                                                                                                                                  HP1033405-A2.
                                                                                                                                                                                                                                                                                                                                                                                                                                             -MAY-1999;
-MAY-1999;
-MAY-1999;
-MAY-1999;
                                                                                                                                     17-OCT-2000
                                                                                                                                                                                                                                                                                             25-MAR-1999;
                                                                                                                                                                                                                                                                                                                                                                          28-APR-1999;
                                                                                                                                                                                                                                                                                                                                                                                         1999;
                                                                                                                                                                                                                                                                                                                                                                                                 MAY-1999;
                                                                                                                                                                                                                                  06-SEP-2000
                                                                                                                                                                                                                                                                                                                                                                                                         MAY-1999;
                                                                                                                                                                                                                                                                                                                                                                                                                06-MAY-1999;
                                                                                                                                                                                                                                                                                                                                                                                                                       06-MAY-1999;
                                                                                                                      AAG16857;
                                                                                                                                                                                                                                                                                                                                                                                 30-APR-3
                                                                                        80
                                                                 275
                                                                                               AAG16857
                                                                                        RESULT
                                 qq
                  ΩŊ
                                                  δŏ
                                                                                                             12;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        || | : :| : | : : | : : | : : | snlvspqiadddddsvsdshiergewsefhvevrrvspyfggstysggskegc----- 141
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            258 VQSDSKRESVCNKADAES -----EPVAQKSQLDRTVCISDAGACGET----LSVTSE 305
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       ENSLVKK-----KERSLSSGSNFCSEQKTSGIINKFCSAKDSEHNEKYEDTFLESEEIGT 360
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              205 SNFTSTHLLLKEDEGVDDVNFRKVRKPKGKVTILKGIPI-----KKTKKGCRKSCSGF 257
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  89;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                4.9%; Score 148.5; DB 21; Length 327; 20.7%; Pred. No. 1.6e-05; tive 47; Mismatches 102; Indels 89;
                                                                                    99US-0151438.
99US-0151930.
99US-0152363.
                                                                                                                                                9905-0155139
9905-0155486
9905-0156458
9905-0156459
9905-0157117
9905-0157753
9905-015785
99US-0149722.
99US-0149723.
99US-0149929.
99US-0149900.
99US-015966.
99US-0151065.
                                                                                                                                                                                                                             99US-0158369.
99US-0159293.
99US-0159294.
                                                                                                                                                                                                                                                                                         990S-0159638
990S-0159584
990S-0160741.
990S-0160767
990S-0160776
990S-0160814
990S-0160815
                                                                                                                                                                                                                                                    99US-0159295.
99US-0159329.
99US-0159330.
                                                                                                                                                                                                                                                                                                                                                                            9905-0161404
9905-0161405
9905-01611405
9905-01611350
9905-01611360
9905-01611361
9905-01611920
9905-01611932
9905-01611932
                                                                                                                                                                                                                                                                          99US-0159331.
                                                                     99US-0151080,
99US-0151303,
                                                                                                           99US-0153070.
                                                                                                                                  99US-0154039.
                                                                                                                           99US-0154018
                                                                                                                                                                                                                                                                                                                                                                       99US-0160989
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                62; Conservative
                                                                                                                                                                                                                                                                                                                                                                                                                                                                        Similarity
              20 - AUG - 1999
23 - AUG - 1999
25 - AUG - 1999
26 - AUG - 1999
27 - AUG - 1999
27 - AUG - 1999
27 - AUG - 1999
31 - AUG - 1999
                                                                                                                                16-SEP-1999,
20-SEP-1999,
23-SEP-1999,
24-SEP-1999,
28-SEP-1999,
04-0-CT-1999,
06-0-CT-1999,
06-0-CT-1999,
                                                                                                                                                                                                             07-0CT-1999;
08-0CT-1999;
12-0CT-1999;
13-0CT-1999;
13-0CT-1999;
14-0CT-1999;
14-0CT-1999;
14-0CT-1999;
14-0CT-1999;
                                                                                                                                                                                                                                                                                                 18-OCT-1999;
21-OCT-1999;
21-OCT-1999;
                                                                                                                                                                                                                                                                                                                                       21-0CT-1999;
21-0CT-1999;
22-0CT-1999;
22-0CT-1999;
22-0CT-1999;
25-0CT-1999;
                                                                                                                                                                                                                                                                                                                                                                                   25-0CT-1999;
25-0CT-1999;
26-0CT-1999;
26-0CT-1999;
26-0CT-1999;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    <u>::</u>
                                                                                                                   13-SEP-1999;
                                                                                                                                                                                                                                                                                                                        21-0CT-1999;
21-0CT-1999;
                                                                                                                                                                                                                                                                                                                                                                                                                                                 29-OCT-1999;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                 Query Match
Best Local 9
                                                                                                                                                                                                                                                                                                                                                                                                                                                                               Matches
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              89
g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           QQ
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             οy
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          ò
```

```
Protein identification; signal transduction pathway; metabolic pathway; hybridisation assay; genetic mapping; gene expression control; promoter; termination sequence.
                                                                                                                                                                                    421 KYNKEALSPPRRKAFKKWTPPRSPFNLVQETLFHDPWKLLIATIFLNRTSGKMA-IPVLW 479
                                                                                                                                                                                                                     195 rgvkvrrvspyfqestvseqpn----qapkglrnyf--------------------- 226
                                                            361 KVEVVERKEHLHTDILKRGSEMDNNCSPTRKDFTGEKIFQEDTIPRTQIERRKTSLYFSS 420
                                                                                                Arabidopsis thaliana protein fragment SEQ ID NO: 17661.
```

15-0138540 15-0138847 15-01384119 15-013945119 15-01394545 15-0139454 15-0139457 15-0139457 15-0139457 15-0139457 15-0139457 15-0139457 15-0139460 15-0139460 15-0139460	990S-0139817. 990S-0140353. 990S-0140353. 990S-0140695. 990S-0141287. 990S-0141287. 990S-0141287. 990S-0142154. 990S-0142154. 990S-0142154. 990S-0142803. 990S-014390. 990S-0143542. 990S-014362. 990S-0144085. 990S-0144085. 990S-0144313. 990S-0144313. 990S-0144313. 990S-0144313. 990S-0144313. 990S-0144314. 990S-0144314. 990S-0144314. 990S-0145086. 990S-0145086. 990S-0145086. 990S-0145087. 990S-0145089. 990S-0145089. 990S-0145089. 990S-0145089. 990S-0145089. 990S-0145089. 990S-0145089.	9US-014703 9US-014720 9US-014730 9US-014716 9US-014726 9US-014741
	21-JUN-1999; 22-JUN-1999; 23-JUN-1999; 24-JUN-1999; 26-JUN-1999; 30-JUN-1999; 30-JUN-1999; 30-JUL-1999; 30-JUL-1999; 30-JUL-1999; 30-JUL-1999; 30-JUL-1999; 30-JUL-1999; 30-JUL-1999; 30-JUL-1999; 30-JUL-1999; 30-JUL-1999; 31-JUL-1999; 31-JUL-1999; 32-JUL-1999; 32-JUL-1999; 32-JUL-1999; 33-JUL-1999; 34-JU	3-AUG-199 4-AUG-199 4-AUG-199 5-AUG-199 5-AUG-199 6-AUG-199
PR PR PR PR PR PR PR	2	7 7 7 7 7 7 7 7 7 7 7 7 7 7 7 7 7 7 7

	÷	
US-0147935 US-0148171 US-0148841 US-01488684 US-0148684 US-0149268 US-0149427 US-014927 US-014927 US-014927 US-014922	990S-0150566. 990S-0150884. 990S-0151065. 990S-0151080. 990S-0151303. 990S-0151330. 990S-0153070. 990S-0153070. 990S-0153070. 990S-0153070. 990S-0153070. 990S-015588. 990S-015588. 990S-015588. 990S-015588.	10.55-01.59.59.59.59.59.59.59.59.59.59.59.59.59.
-Aug-1999 -Aug-1999 -Aug-1999 -Aug-1999 -Aug-1999 -Aug-1999 -Aug-1999 -Aug-1999 -Aug-1999	25-AUG-1999; 26-AUG-1999; 27-AUG-1999; 27-AUG-1999; 31-AUG-1999; 31-AUG-1999; 31-AUG-1999; 10-SEP-1999; 10-SEP-1999; 15-SEP-1999; 16-SEP-1999; 20-SEP-1999; 20-SEP-1999; 20-SEP-1999; 20-CCT-1999;	1-007-199 1-007-199 1-007-199 1-007-199 1-007-199 1-007-199 1-007-199 1-007-199 1-007-199 1-007-199 1-007-199 1-007-199 1-007-199 1-007-199 1-007-199 1-007-199 1-007-199 1-007-199 1-007-199
9 8 8 8 8 8 8 8 8 8 8 8 8 8 8 8 8 8 8 8		7

Ouery Match 4.9%; Score 148.5; DB 21; Length 352; Best Local Similarity 20.7%; Pred. No. 1.9e-05; Matches 62; Conservative 47; Mismatches 102; Indels 89; Gaps 12;

QY 205 SNFTSTHLLLKEDEGVDDVNFRKVRKPKGKVTILKGIPI-----KKTKKGCRKSCSGF 257

```
The present sequence represents the protein encoded by exon 11 of the BRCA2 cancer susceptibility gene. The nucleic acid molecule can be used
258 VQSDSKRESVCNKADAES-----EPVAQKSQLDRTVCISDAGACGET----LSVTSE 305
                                                                    167 ---ds--dsvcskegcskvqakvprvspyfqastisq--cdsdivsssqsgrnyrkgssk 219
                                                                                              306 ENSLVKK-----KERSLSSGSNFCSEQKTSGIINKFCSAKDSEHNEKYEDTFLESEEIGT 360
                                                                                                                                                     361 KVEVVERKEHLHTDILKRGSEMDNNCSPTRKDFTGEKIFQEDTIPRTQIERRKTSLYFSS 420
                                                                                                                                                                               252 --kvvkvsryfhadgiqvnesqkeksrnvrk-----tpivspvlslsgktddvyl-- 299
                                                                                                                                                                                                            421 KYNKEALSPPRRKAFKKWTPPRSPFNLVQETLFHDPWKLLIATIFLNRTSGKMA-IPVLW 479
                                                                                                                                                                                                                             Nucleic acid molecules comprising part or all of the BRCA2 cancer susceptibility gene · useful for diagnosis, prognosis or therapeutic
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     /note= "Encoded by the end of exon 11 and the start
  of exon 12"
                                                                                                                                                                                                                                                                                                                                                                                                                      BRCA2 cancer susceptibility gene; breast cancer; ovarian cancer; gene therapy; prostate cancer; colorectal cancer; ocular melanoma; leukaemia; human.
                                                                                                                                                                                                                                                                                                                                                                                           BRCA2 cancer susceptibility gene protein encoded by exon 11.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              Stratton MR, Wooster RF;
                                                                                                                         220 rqvkvrrvspyfqestvseqpn----qapkglrnyf----
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            /note= "Encoded by GNA"
Misc-difference 1626
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    'note= "Encoded by GRA"
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         'note= "Encoded by GTN"
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   (CANC-) CANCER RES CAMPAIGN TECHNOLOGY
                                                                                                                                                                                                                                                                                                           AAW25049 standard; Protein; 1644 AA.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           Key Location/Qualifiers
Misc-difference 81
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     Claim 10; Fig 7; 124pp; English.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    96GB-0024453.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            96СВ-0017961.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              95GB-0023959
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        95GB-0025555
                                                                                                                                                                                                                                                                                                                                                                 03-FEB-1998 (first entry)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                Futreal PA,
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                Misc-difference 1624
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       Misc-difference 1644
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      WPI; 1997-261854/24.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             treatment of cancer
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   (UYDU-) UNIV DUKE.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        N-PSDB; AAT79680
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   25-NOV-1996;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              28-AUG-1996;
23-NOV-1995;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              GB2307477-A.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       14-DEC-1995;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             Ashworth A,
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        28-MAY-1997
                                                                                                                                                                                                                                                                                                                                       AAW25049;
                                         ŏ
                                                                    qq
                                                                                                ò
                                                                                                                         qq
                                                                                                                                                       δ
                                                                                                                                                                                Db
                                                                                                                                                                                                            ò
                                                                                                                                                                                                                                       pp
```

```
23;
                                                                                                                                                                                                                                                                                                                                                                              So that the expression of the BRCA2 nucleic acid is reduced or prevented. The nucleic acid molecules are also useful in a method for diagnosing susceptibility or predsposition to cancer in a patient. The nucleic acid molecules are used to design probes or primers for PCR to determine or detect the presence of mutations in a sample of nucleic acid molecules are used to design probes or primers for PCR to acid from a patient. The BRCA2 promoter region is useful for screening for substances which modulate the expression of nucleic acid under control to the promoter. Antibodies are used to determine the presence, The polypeptides are used to screen for binding partners, these are used to screen for binding partners; these are useful to screen for substances which minic the activity of BRCA2 are
                                                                                  sequence. The BRCA2 nucleic acid molecules and proteins are useful in a method of medical treatment, preferably gene therapy, especially for treating cancer, where the cancer is female or male breast cancer, ovarian, prostate or colorectal cancer, ocular melanoma or leukaemia. In particular antisense oligonucleotides capable of hybridising to the BRCA2 nucleic acid, pre-mRNA or mature mRNA are used
to construct probes for screening cDNA or genomic libraries, sequencing positive clones obtained, and assembling the full length BRCA2
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            538 qvdsskqfegtvelkrkfagllkndcnksasgyltde-----nevgfrgfysahgtkl 590
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   155 YKDCSMAALITSHL--QNQSNNSNWNLRTRSKCKKDVFMPPSSSSELQESRGLSNFTSTHL 212
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        213 LLKEDEGVDDVNFRKVRK------239
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       335 ksdisluidkipeknndymnkwag----llgpisnhsfggsfrtasnkeiklsehnikk 389
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         96 QRLFGKTAGRFDVYFISPQGLKFRSKSSLANYLHKNGETSL-KPEDFDFTVLSKRGIKSR 154
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           441 vsdcknshitpqmlfskqdfnsnhnl-----tpsqkaeite-----lstil 481
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           285 D-RTVCISDAGACGETLSVTSEENSLVKKKERSLSSGSNFCSEQKTSGIINKFCSAKDSE 343
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   591 nvstealgkavklfsdieniseets---aevhpislssskchd----svvsmf---kien 640
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           344 HNEKYEDTFLESEE-----IGTKVEVVERKEHLHTDILKRGSEMDNN--CSP 388
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   36 KEDVAMELERVGEDEEQMMIKRSSECNPLLQEPIASAQFGATAGTECRKSVPCGWERVVK 95
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            240 -----GIPIKK----TKKGCRKSCSGFVQSDSKRESVCNKADAESEPVAQKSQL
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           Ouery Match 4.7%; Score 145; DB 18; Length 1644; Best Local Similarity 20.9%; Pred. No. 0.00051; Matches 96; Conservative 67; Mismatches 158; Indels 138.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         Human; breast cancer; susceptibility; gene; BRCA2; diagnosis;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             polypeptide, which can be used as cancer therapeutics.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 Human breast cancer susceptibility gene BRCA2 product.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  Conservative 67; Mismatches 158;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   :| :| |: :| |: ::| |: ::| |: ::| |: ::| |: ::| |: ::| |: ::| |: ::| |: ::| |: ::| |: ::| |: ::| |: ::| |: ::| |: ::| |: ::| |: ::| |: ::| |: ::| |: ::| |: ::| |: ::| |: ::| |: ::| |: ::| |: ::| |: ::| |: ::| |: ::| |: ::| |: ::| |: ::| |: ::| |: ::| |: ::| |: ::| |: ::| |: ::| |: ::| |: ::| |: ::| |: ::| |: ::| |: ::| |: ::| |: ::| |: ::| |: ::| |: ::| |: ::| |: ::| |: ::| |: ::| |: ::| |: ::| |: ::| |: ::| |: ::| |: ::| |: ::| |: ::| |: ::| |: ::| |: ::| |: ::| |: ::| |: ::| |: ::| |: ::| |: ::| |: ::| |: ::| |: ::| |: ::| |: ::| |: ::| |: ::| |: ::| |: ::| |: ::| |: ::| |: ::| |: ::| |: ::| |: ::| |: ::| |: ::| |: ::| |: ::| |: ::| |: ::| |: ::| |: ::| |: ::| |: ::| |: ::| |: ::| |: ::| |: ::| |: ::| |: ::| |: ::| |: ::| |: ::| |: ::| |: ::| |: ::| |: ::| |: ::| |: ::| |: ::| |: ::| |: ::| |: ::| |: ::| |: ::| |: ::| |: ::| |: ::| |: ::| |: ::| |: ::| |: ::| |: ::| |: ::| |: ::| |: ::| |: ::| |: ::| |: ::| |: ::| |: ::| |: ::| |: ::| |: ::| |: ::| |: ::| |: ::| |: ::| |: ::| |: ::| |: ::| |: ::| |: ::| |: ::| |: ::| |: ::| |: ::| |: ::| |: ::| |: ::| |: ::| |: ::| |: ::| |: ::| |: ::| |: ::| |: ::| |: ::| |: ::| |: ::| |: ::| |: ::| |: ::| |: ::| |: ::| |: ::| |: ::| |: ::| |: ::| |: ::| |: ::| |: ::| |: ::| |: ::| |: ::| |: ::| |: ::| |: ::| |: ::| |: ::| |: ::| |: ::| |: ::| |: ::| |: ::| |: ::| |: ::| |: ::| |: ::| |: ::| |: ::| |: ::| |: ::| |: ::| |: ::| |: ::| |: ::| |: ::| |: ::| |: ::| |: ::| |: ::| |: ::| |: ::| |: ::| |: ::| |: ::| |: ::| |: ::| |: ::| |: ::| |: ::| |: ::| |: ::| |: ::| |: ::| |: ::| |: ::| |: ::| |: ::| |: ::| |: ::| |: ::| |: ::| |: ::| |: ::| |: ::| |: ::| |: ::| |: ::| |: ::| |: ::| |: ::| |: ::| |: ::| |: ::| |: ::| |: ::| |: ::| |: ::| |: ::| |: ::| |: ::| |: ::| |: ::| |: ::| |: ::| |: ::| |: ::| |: ::| |: ::| |: ::| |: ::| |: ::| |: ::| |: ::| |: ::| |: ::| |: ::| |: ::| |: ::| |: ::| |: ::| |: ::| |: ::| |: ::| |: ::| |: ::| |: ::| |: ::| |: ::| |: ::| |: ::| |: ::| |: ::| |: ::| |: ::| |: ::| |: ::| |: ::| |: ::| |: ::| |: ::| |: ::| |: ::| |: ::| |: ::| |: ::| |: ::| |: :
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          389 TRK----DFTGEKIFQEDTIPRTQIERRKTSLYFSSKYN 423
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          AAW19211 standard; Protein; 3418 AA
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 screening; treatment; gene therapy.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         (first entry)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    1644 AA;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    Sequence
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       AAW19211:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             RESULT 10
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           AAW19211
88.56666666666666668888
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               ōγ
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 QD
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         22
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   QΫ́
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           20
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             δ
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     qq
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            οŽ
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             q
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               οy
```

```
389 TRK----DFTGEKIFQEDTIPRTQIERRKTSLYFSSKYN 423
                                                                                     RESULT
                                                                                                                                                                                  ŏ
                                  qq
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           23;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               1078 vsdcknshitpgmlfskqdfnsnhnl-----tpsgkaeite----lstil 1118
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  972 ksdislnidkipeknndymnkwag----llgpisnhsfggsfrtasnkeiklsehnikk 1026
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            nvstealqkavklfsdieniseets---aevhpislssskchd----svvsmf---kien 1277
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            1119 ----eesgsqfeftqfrkpsyilqkstfevpenqmtilkttseecrdadlhvimnapsig 1174
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             285 D-RTVCISDAGACGETLSVTSEENSLVKKKERSLSSGSNFCSEQKTSGIINKFCSAKDSE 343
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            155 YKDCSMAALTSHL--QNQSNNSNWNLRTRSKCKKDVFMPPSSSSELQESRGLSNFTSTHL 212
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               HNEKYEDTFLESEE------IGTKVEVVERKEHLHTDILKRGSEMDNN--CSP 388
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        Query Match
4.7%; Score 145; DB 18; Length 3418;
Best Local Similarity 20.9%; Pred. No. 0.0017;
Matches 96; Conservative 67; Mismatches 158; Indels 138; Gaps
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          QRLFGKTAGRFDVYFISPQGLKFRSKSSLANYLHKNGETSL-KPEDFDFTVLSKRGIKSR 154
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            36 KEDVAMELERVGEDEEQMMIKRSSECNPLLQEPIASAQFGATAGTECRKSVPCGWERVVK 95
                                                                                                                                                                                                                                                                                                                                                                                                                                     diagnosing breast cancer and screening for compounds to treat breast
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     The present sequence is the human breast cancer susceptibility gene BRCA2 product, which can be used to diagnose breast cancer and screen for compounds to treat breast cancer. BRCA2 can also be used in gene therapy to restore wild type BRCA2 gene function to a cell, which has lost its or has altered (i.e. by virtue of a mutation in
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               -----GIPIKK-----TKKGCRKSCSGFVQSDSKRESVCNKADAESEPVAQKSQL
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  Human breast cancer susceptibility gene BRCA2 - useful for
                                                                                                                                                                                                                                                                                                                        Tavtigian SV;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              213 LLKEDEGVDDVNFRKVRK-------PKGKVTILK------
                                                                                                                                                                                                                                                                                                                        Kamb A, Rommens JM, Simard J,
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         Claim 1; Pages 90-106; 189pp; English.
                                                                                                                                                     95US-0573779.
95US-0575359.
95US-0576559.
96US-0585391.
                                                                                                    96WO-US19598
                                                                                                                                     96US-0639501
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          BRCA2) BRCA2 gene function
                                                                                                                                                                                                                                                    MYRIAD GENETICS I.
UNIV PENNSYLVANIA.
                                                                                                                                                                                                                                      DEV LP.
                                                                                                                                                                                                                                                                                       (ENDO-) ENDO RECH INC
                                                                                                                                                                                                                                                                                                                                                                       WPI; 1997-341680/31.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           3418 AA;
                                                                                                                                                                                                                                                                                                                                                                                        N-PSDB; AAT69707
                                                                                                                                                                                                                                      HSCR-) HSC RES
                                                                                                                                                                 20-DEC-1995;
21-DEC-1995;
11-JAN-1996;
      Homo sapiens.
                                      WO9722689-A1
                                                                                                    17-DEC-1996;
                                                                                                                                     29-APR-1996;
                                                                                                                                                      18-DEC-1995
                                                                     26-JUN-1997
                                                                                                                                                                                                                                                                                                                                      Weber BL;
                                                                                                                                                                                                                                                                                                                        Couch F,
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             Sequence
                                                                                                                                                                                                                                                        MYRI-)
                                                                                                                                                                                                                                                                       (UYPE-)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              96
```

g

ò

ò

ò

Ω

ò

ò

```
1027 skmffk----dieegyptsl---acveivntlaldnqkkiskpqsin-tvsahlqssvv 1077
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               972 ksdisInidkipeknndymnkwag-----llgpisnhsfggsfrtasnkeiklsehnikk 1026
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  96 QRLFGKTAGRFDVYFISPQGLKFRSKSSLANYLHKNGETSL-KPEDFDFTVLSKRGIKSR 154
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            4.7%; Score 145; DB 18; Length 3418;
20.9%; Pred. No. 0.0017;
ive 67; Mismatches 158; Indels 138; Gaps
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       that inhibits the growth of breat and overian suppressor, the contains a granin, region, indicating that it is a secreted proteins a proteins. BRCA2 protein and DNA sequences encoding BRCA2 protein and DNA sequences encoding BRCA2 (see AAT84841) can be used in the treatment of breast and ovarian cancers. Since the BRCA2 protein is secreted, it can be used to identify the BRCA2 receptor and hence to identify BRCA2 protein mimetics which act on the receptor for use in breast and ovarian cancer treatment. The BRCA1 gene (see AAT84840) and BRCA2 protein (see AAW23286) have also been characetised. Methods are claimed for the isolation of BRCA1 or BRCA2 receptors, and for treating or preventing breast and (gene-linked hereditary or sporadic) ovarian cancers using BRCA1 and BRCA2 proteins and genes.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         36 KEDVAMELERVGEDEEQMMIKRSSECNPLLQEPIASAQFGATAGTECRKSVPCGWERVVK 95
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        inhibit
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    BRCA1 and BRCA2 tumour suppressor gene products - useful to inhibit breast and ovarian cancer cell growth and tumourigenesis, or treat gene linked hereditary or sporadic ovarian or breast cancer
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    This protein comprises the human BRCA2 tumour suppressor protein
                                                                                                                                                                                                                                                                                                                                                                                                                    Human breast and ovarian cancer susceptibility protein BRCA2.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        Page DL;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            tumour suppressor protein; gene therapy; receptor.
:| :| | ::| | 1328 srnshnlefdqsdsskndtv---cihkdetdllftdqhm 1363
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        Jensen RA, Jetton TL,
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           BRCA2; breast cancer; ovarian cancer; human;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          Thompson ME;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  Claim 1.8; Page 72-88; 148pp; English.
                                                                                                                                                                                                        AAW23287 standard; Protein; 3418 AA
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             Szabo CI,
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               97WO-US03340.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          96us-0603753.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    20.9%;
                                                                                                                                                                                                                                                                                                                                                  21-FEB-1998 (first entry)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        Best_Local Similarity 20.99
Matches 96; Conservative
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            (UYVA-) UNIV VANDERBILT.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     (UNIW ) UNIV WASHINGTON
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            Holt JT
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             Robinson-Benion CL,
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                WPI; 1997-434733/40.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               3418 AA;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  N-PSDB; AAT84841
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            Claire-King M,
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       Homo sapiens.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             20-FEB-1996;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   19-FEB-1997;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          WO9730108-A1
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  Sequence
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       Query Match
                                                                                                                                                                                                                                                                               AAW23287;
                                                                                                                                                                          AAW23287
```

```
The present invention describes genomic DNA which contains a BRCA2 gene
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    where the first 12 nucleotides beginning exon 5 are 5'-TCTGTTGTTGT-3' as in sequence (I) (see AAX03249), where nucleotides numbers 5782-5790 are GTTGGTT as in sequence (IV) (see AAX3025), and where the last 20 nucleotides encoding exon 15 are 5'-CTGGGTGTTCTGATAAACAG-3' as in sequence (II) (see AAX30251) and the first 20 nucleotides beginning exon 16 are 5'-CTGTATACGTATGCGGTTTC-3' as in sequence (III)
                            1078 vsdcknshitpqmlfskqdfnsnhnl-----tpsqkaeite-----lstil 1118
                                                                                       1175 qvdsskqfegtveikrkfagllkndcnksasgyltde-----nevgfrgfysahgtkl 1227
                                                                                                                                                                                                                                                                                                    1278 hndk---tvseknnkcqlilgnniemttgtfveei-----tenykrntenednkytaa 1327
                                                                                                                                                                                                                                   1228 nvstealqkavklfsdieniseets---aevhpislssskchd----svvsmf---kien 1277
155 YKDCSMAALTSHL--QNQSNNSNWNLRTRSKCKKDVFMPPSSSSELQESRGLSNFTSTHL 212
                                                                                                                                                                                                   285 D-RTVCISDAGACGETLSVTSEENSLVKKKERSLSSGSNFCSEQKTSGIINKFCSAKDSE 343
                                                                                                                                  240 ------GIPIKK----TKKGCRKSCSGFVQSDSKRESVCNKADAESEPVAQKSQL 284
                                                                                                                                                                                                                                                                     344 HNEKYEDTFLESEE------IGTKVEVVERKEHLHTDILKRGSEMDNN--CSP 388
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         Human; BRCA2; genetic testing; protein therapy; haplotype; detection; gene therapy; breast cancer; ovarian cancer.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    New coding sequence haplotypes of the human BRCA2 gene - used to develop products for determining susceptibility to, detection and
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   Jackson GM, Mürphy PD, Olson SJ, Park M;
hryer B, White MB, Yoshikawa M;
                                                                                                                                                                                                                                                                                                                                                         389 TRK----DFTGEKIFQEDTIPRTQIERRKTSLYFSSKYN 423
                                                                     Claim 19; Page 112-118; 226pp; English.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    treatment of breast or ovarian cancer
                                                                                                                                                                                                                                                                                                                                                                                                                                                        AAY04356 standard; Protein; 3418 AA
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  97US-0055784.
97US-0064926.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              98WO-US16905.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              98US-0084471.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  98US-0071715.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           Human BRCA2 (omi3) protein.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               97US-0065367
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           21-JUN-1999 (first entry)
                                                                  213 LLKEDEGVDDVNFRKVRK-
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 Schryer B,
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   (ONCO-) ONCORMED INC
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 WPI; 1999-190163/16.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     N-PSDB; AAX30257
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            WO9909164-A1.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             Homo sapiens
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            14-AUG-1998;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              22-MAY-1998;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 Eskandari T,
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            25-FEB-1999.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             07-NOV-1997
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 01-MAY-1998;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                -AUG-1997
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     Rabin MB,
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          AAY04356;
                                                                                                                                                                                                                                                                                                                                                                                                                          RESULT 12
                                                                                                                                                                                                                                                                                                                                                                                                                                          AAY04356
                                                                ó
                                                                                                                                                                 QQ
                                                                                                                                                                                                                                                                                                    qq
                                                                                                                                                                                                                                                                                                                                                                    qq
                                                                                                                                                                                                   ò
                                                                                                                                                                                                                                                                   δy
```

```
(see AAX30253). Products and methods from the present invention can be used for identifying mutations in the BRCA2 gene leading to predisposition or higher susceptibility to breast or ovarian cancer. They can also be used for detection and gene therapy for breast and ovarian cancers. They can be used in methods for monitoring disease progression, for determining patients suited for gene and protein replacement progression, or for detecting the presence or quantifying the amount of a pumper proxych inhibitor following such therapy. The BRCA2 protein, polypeptides, their functional equivalents, antibodies, and PNs may also be useful in the study of the characteristics of BRCA2 proteins, such as structure and function of BRCA2 protein in normal and cancerous cells. The present sequence represents human BRCA2 (omi3).
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                23;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      972 ksdislnidkipeknndymdkwag----llgpisnhsfggsfrtasnkeiklsehnikk 1026
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            | : | | : | | : | | : | | : | | : | | : | | : | | : | | : | | : | | : | | : | | : | | : | | : | | : | | : | | : | | : | | : | | : | | : | | : | | : | | : | | : | | : | | : | | : | | : | | : | | : | | : | | : | | : | : | | : | : | | : | : | | : | : | | : | : | | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | 
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           1078 vsdcknshitpqmlfskqdfnsnhnl------tpsqkaeite----1stil 1118
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     1119 ----eesgsqfeftqfrkpsyilqkstfevpengmtilkttseecrdadlhvimnapsig 1174
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                1228 nvstealqkavklfsdieniseets---aevhpislssskchd----svvsmf---kien 1277
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     1175 gvdsskgfegtveikrkfagllkndcnksasgyltde-----nevgfrgfysahgtkl 1227
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          278 hndk---tvseknnkcqlilqnniemttgtfveei-----tenykrntenednkytaa 1327
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                155 YKDCSMAALTSHL--QNQSNNSNWNLRTRSKCKKDVFWPPSSSSELQESRGLSNFTSTHL 212
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           96 QRLFGKTAGRFDVYFISPQGLKFRSKSSLANYLHKNGETSL-KPEDFDFTVLSKRGIKSR 154
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                240 -----GIPIKK----TKKGCRKSCSGFVQSDSKRESVCNKADAESEPVAQKSQL 284
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     285 D-RTVCISDAGACGETLSVTSEENSLVKKKERSLSSGSNFCSEQKTSGIINKFCSAKDSE 343
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               344 HNEKYEDTFLESEE------IGTKVEVVERKEHLHTDILKRGSEMDNN--CSP 388
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     Gaps
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            Human; BRCA2; genetic testing; protein therapy; haplotype; detection;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      36 KEDVAMELERVGEDEEQMMIKRSSECNPLLQEPIASAQFGATAGTECRKSVPCGWERVVK 95
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          213 LLKEDEGVDDVNFRKVRK-------PKGKVTILK--------
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      67; Mismatches 158; Indels 138;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           4.7%; Score 145; DB 20; Length 3418; 20.9%; Pred. No. 0.0017;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      :| ::| | ::| | ::| | 3328 srnshnlefdgsdsskndtv---cihkdetdllftdqhn 1363
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          389 TRK----DFTGEKIFQEDTIPRTQIERRKTSLYFSSKYN 423
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   breast cancer; ovarian cancer.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            AAY04357 standard; Protein; 3418 AA.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 Human BRCA2 (omi4) protein.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      96; Conservative
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 21-JUN-1999 (first entry)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   3418 AA;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         Local Similarity
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              qene therapy;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              Homo sapiens
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   25-FEB-1999.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        Sednence
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            AAY04357;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                Query Match
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 RESULT 13
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      Matches
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    AAY04357
               0000000000000000×3
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           qq
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                q
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     qq
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           qq
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     Ωÿ
```

```
23;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     tumour growth inhibitor following such therapy. The BRCA2 protein, polypeptides, their functional equivalents, antibodies, and PNs may also be useful in the study of the characteristics of BRCA2 proteins, such as structure and function of BRCA2 in oncogenesis or subcellular localisation of BRCA2 protein in normal and cancerous cells. The present sequence represents human BRCA2 (omi4).
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          1078 vsdcknshitpqmlfskqdfnsnhnl-----tpsqkaeite----1stil 1118
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             1228 nvstealgkavklfsdieniseets---aevhpislssskchd----svvsmf---kien 1277
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               972 ksdislnidkipeknndymnkwag----llgpisnhsfggsfrtasnkeiklsehnikk 1026
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           1119 ----eesgsgfeftgfrkpsyilgkstfevpengmtilkttseecrdadlhvimnapsig 1174
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        1175 qvdsskqfegtveikrkfagllkndcnksasgyltde-----nevgfrgfysahgtkl 1227
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                155 YKDCSMAALTSHL--QNQSNNSNWNLRTRSKCKKDVFMPPSSSSELQESRGLSNFTSTHL 212
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           96 ORLFGKTAGREDVYFISPQGLKFRSKSSLANYLHKNGETSL-KPEDFDFTVLSKRGIKSR 154
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    285 D-RTVCISDAGACGETLSVTSEENSLVKKKERSLSSGSNFCSEQKTSGIINKFCSAKDSE 343
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          Gaps
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   36 KEDVAMELERVGEDEEQMMIKRSSECNPLLQEPIASAQFGATAGTECRKSVPCGWERVVK 95
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               -----GIPIKK----TKKGCRKSCSGFVQSDSKRESVCNKADAESEPVAQKSQL
                                                                                                                                                                                                                    New coding sequence haplotypes of the human BRCA2 gene - used to develop products for determining susceptibility to, detection and
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         67; Mismatches 158; Indels 138;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              Local Similarity 20.9%; Pred. No. 0.0017; Los 96; Conservative
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         213 LLKEDEGVDDVNFRKVRK-------PKGKVTILK------
                                                                                                                                      Olson SJ, Park M;
                                                                                                                                      Jackson GM, Murphy PD, Olson S.
hrver B, White MB, Yoshikawa M;
                                                                                                                                                                                                                                                                               Claim 29; Page 133-140; 226pp; English.
                                                                                                                                                                                                                                                    treatment of breast or ovarian cancer
                                                                                                                                                    White MB,
                        98US-0084471.
97US-0055784.
97US-0064926.
97US-0065367.
98US-0071715.
98WO-US16905
                                                                                                                                                     Schryer B,
                                                                                                               (ONCO-) ONCORMED INC.
                                                                                                                                                                                 WPI; 1999-190163/16.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        3418 AA;
                                                                                                                                                                                              N-PSDB; AAX30258
                                                    07-NOV-1997;
12-NOV-1997;
01-MAY-1998;
14-AUG-1998;
                           22-MAY-1998;
                                                                                                                                        Eskandari T,
                                          15-AUG-1997
                                                                                                                                                        Rabin MB,
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           Sequence
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   Query Match
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  240
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             Matches
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        a
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 ò
  ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      ò
```

```
The present invention describes genomic DNA which contains a BRCA2 gene where the first 12 nucleotides beginning exon 5 are 5' TCCTGTTGTTGTT3' cas in sequence (I) (see AAX302149), where nucleotides numbers cas in sequence (I) (see AAX30251), and where the last 20 nucleotides encoding exon 15 are 5'-CTGCTGTTGTCATAAACAG-3' cas in sequence (II) (see AAX30251) and the first 20 nucleotides compared (II) (see AAX30251), and the first 20 nucleotides compared (III) (see AAX30251). Products and methods from the present invention can be used for identifying mutations in the BRCA2 gene leading to predisposition or higher susceptibility to breast or ovarian cancer. They can also be used for detection and gene therapy for breast and ovarian cancers. They can be used in methods for monitoring disease progression, or for detecting the presence or quantifying the amount of a trongression, or for detecting the presence or quantifying the amount of compared to the characteristics of BRCA2 proteins, such as the useful in the study of the characteristics of BRCA2 proteins, such as continuous feature and function of BRCA2 in oncogenesis or subcellular repersent
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            localisation of BRCA2 protein in normal and cancerous cells. The present
                   Human; BRCA2; genetic testing; protein therapy; haplotype; detection; gene therapy; breast cancer; ovarian cancer.
----IGTKVEVVERKEHLHTDILKRGSEMDNN--CSP 388
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         New coding sequence haplotypes of the human BRCA2 gene - used to develop products for determining susceptibility to, detection and
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                Park M;
                                                                                           Eskandari T, Jackson GM, Murphy PD, Olson SJ,
Rabin MB, Schryer B, White MB, Yoshikawa M;
                                                                      389 TRK----DFTGEKIFQEDTIPRTQIERRKTSLYFSSKYN 423
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 sequence represents human BRCA2 (omi5).
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 Claim 20; Page 155-162; 226pp; English
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               treatment of breast or ovarian cancer
                                                                                                                                                                                                      AAY04358 standard; Protein; 3418 AA
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   98WO-US16905.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     98US-0084471.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        97US-0055784.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           97US-0064926.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        97US-0065367.
98US-0071715.
                                                                                                                                                                                                                                                                                                                Human BRCA2 (omi5) protein.
                                                                                                                                                                                                                                                                            (first entry)
 344 HNEKYEDTFLESEE----
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                (ONCO-) ONCORMED INC
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         WPI; 1999-190163/16.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          N-PSDB; AAX30259
                                                                                                                                                                                                                                                                                                                                                                                                            Homo sapiens.
                                                                                                                                                                                                                                                                                                                                                                                                                                            W09909164-A1.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   14-AUG-1998;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       22-MAY-1998;
15-AUG-1997;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           07-NOV-1997;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           01-MAY-1998;
                                                                                                                                                                                                                                                                            21-JUN-1999
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          12-NOV-1997,
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                25-FEB-1999
                                                                                                                                                                                                                                          AAY04358;
                                                                                                                                                                                    AAY04358
                                                                                                                                                                                                                                                                                             qq
                                                                              ó
```

3418 AA;

Seguence

```
23;
                                                                                                                                                                1027 skmffk----dieeqypts1---acveivntlaldngkklskpgsin-tvsahlgssvv 1077
                                                                                                                                                                                                                                                                                                                                                                                                                            1078 vsdcknshitpqmlfskqdfnsnhnl-----tpsqkaeite-----lstil 1118
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      1119 ...-eesgsqfeftqfrkpsyilqkstfevpenqmtilkttseecrdadlhvimnapsig 1174
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               1175 gvdsskqfegtveikrkfagllkndcnksasgyltde-----nevgfrgfysahgtkl 1227
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     1228 nvstealqkavklfsdieniseets---aevhpislssskchd----svvsmf---kien 1277
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              155 YKDCSMAALTSHL--QNQSNNSNWNLRTRSKCKKDVFMPPSSSSELQESRGLSNFTSTHL 212
                                                                                                                                                                                                                                           96 QRLFGKTAGRFDVYFISPQGLKFRSKSSLANYLHKNGETSL-KPEDFDFTVLSKRGIKSR 154
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    240 ------GIPIKK----TKKGCRKSCSGFVQSDSKRESVCNKADAESEPVAQKSQL 284
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             285 D-RTVCISDAGACGETLSVTSEENSLVKKKERSLSSGSNFCSEQKTSGIINKFCSAKDSE 343
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   344 HNEKYEDTFLESEE-----IGTKVEVVERKEHLHTDILKRGSEMDNN--CSP 388
                                                                    Indels 138; Gaps
                                                                                                                  36 KEDVAMELERVGEDEEQMMIKRSSECNPLLQEPIASAQFGATAGTECRKSVPCGWERVVK 95
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    Human; BRCA2; genetic testing; protein therapy; haplotype; detection; gene therapy; breast cancer; ovarian cancer.
4.7%; Score 145; DB 20; Length 3418; 20.9%; Pred. No. 0.0017; ative 67; Mismatches 158; Indels 138
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        Olson SJ, Park M;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              213 LLKEDEGVDDVNFRKVRK-------PKGKVTILK-----
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    :| :| | : :| |: :| |: :| |: :| |: :| |: :| |: ::| |: ::| |: ::| |: ::| |: ::| |: ::| |: ::| |: ::| |: ::| |: ::| |: ::| |: ::| |: ::| |: ::| |: ::| |: ::| |: ::| |: ::| |: ::| |: ::| |: ::| |: ::| |: ::| |: ::| |: ::| |: ::| |: ::| |: ::| |: ::| |: ::| |: ::| |: ::| |: ::| |: ::| |: ::| |: ::| |: ::| |: ::| |: ::| |: ::| |: ::| |: ::| |: ::| |: ::| |: ::| |: ::| |: ::| |: ::| |: ::| |: ::| |: ::| |: ::| |: ::| |: ::| |: ::| |: ::| |: ::| |: ::| |: ::| |: ::| |: ::| |: ::| |: ::| |: ::| |: ::| |: ::| |: ::| |: ::| |: ::| |: ::| |: ::| |: ::| |: ::| |: ::| |: ::| |: ::| |: ::| |: ::| |: ::| |: ::| |: ::| |: ::| |: ::| |: ::| |: ::| |: ::| |: ::| |: ::| |: ::| |: ::| |: ::| |: ::| |: ::| |: ::| |: ::| |: ::| |: ::| |: ::| |: ::| |: ::| |: ::| |: ::| |: ::| |: ::| |: ::| |: ::| |: ::| |: ::| |: ::| |: ::| |: ::| |: ::| |: ::| |: ::| |: ::| |: ::| |: ::| |: ::| |: ::| |: ::| |: ::| |: ::| |: ::| |: ::| |: ::| |: ::| |: ::| |: ::| |: ::| |: ::| |: ::| |: ::| |: ::| |: ::| |: ::| |: ::| |: ::| |: ::| |: ::| |: ::| |: ::| |: ::| |: ::| |: ::| |: ::| |: ::| |: ::| |: ::| |: ::| |: ::| |: ::| |: ::| |: ::| |: ::| |: ::| |: ::| |: ::| |: ::| |: ::| |: ::| |: ::| |: ::| |: ::| |: ::| |: ::| |: ::| |: ::| |: ::| |: ::| |: ::| |: ::| |: ::| |: ::| |: ::| |: ::| |: ::| |: ::| |: ::| |: ::| |: ::| |: ::| |: ::| |: ::| |: ::| |: ::| |: ::| |: ::| |: ::| |: ::| |: ::| |: ::| |: ::| |: ::| |: ::| |: ::| |: ::| |: ::| |: ::| |: ::| |: ::| |: ::| |: ::| |: ::| |: ::| |: ::| |: ::| |: ::| |: ::| |: ::| |: ::| |: ::| |: ::| |: ::| |: ::| |: ::| |: ::| |: ::| |: ::| |: ::| |: ::| |: ::| |: ::| |: ::| |: ::| |: ::| |: ::| |: ::| |: ::| |: ::| |: ::| |: ::| |: ::| |: ::| |: ::| |: ::| |: ::| |: ::| |: ::| |: ::| |: ::| |: ::| |: ::| |: ::| |: ::| |: ::| |: ::| |: ::| |: ::| |: ::| |: ::| |: ::| |: ::| |: ::| |: ::| |: ::| |: ::| |: ::| |: ::| |: ::| |: ::| |: ::| |: ::| |: ::| |: ::| |: ::| |: ::| |: ::| |: ::| |: ::| |: ::| |: ::| |: ::| |: ::| |: ::| |: ::| |: ::| |: ::| |: ::| |: ::| |: ::| |: ::| |: ::| |: ::| |: ::| |: ::| |: ::| |
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                389 TRK----DFTGEKIFQEDTIPRTQIERRKTSLYFSSKYN 423
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    Jackson GM, Murphy PD, Olson SC
hrver B, White MB, Yoshikawa M;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           AAY04354 standard; Protein; 3418 AA.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        98US-0084471.
97US-0055784.
97US-0064926.
97US-0065367.
98US-0071715.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    98WO-US16905.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     Human BRCA2 (omil) protein.
                                                              Conservative
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    21-JUN-1999 (first entry)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               Schryer B,
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   (ONCO-) ONCORMED INC.
                             Similarity
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   WPI; 1999-190163/16.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    N-PSDB; AAX30255.
      Local Sim.
96;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            Homo sapiens.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          WO9909164-A1.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       Eskandari T,
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                14-AUG-1998;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               22-MAY-1998;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     25-FEB-1999,
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     01-MAY-1998;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               15-AUG-1997
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          07-NOV-1997
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  Rabin MB,
Query Match
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            AAY04354;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      15
                                                          Matches
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      RESULT 1
                                                                                                                                                                                pp
                                                                                                                                                                                                                                               Óχ
                                                                                                                                                                                                                                                                                                      g
                                                                                                                                                                                                                                                                                                                                                                 δ
                                                                                                                                                                                                                                                                                                                                                                                                                            g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              ŏ
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    ŏ
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        q
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      ŏ
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       qq
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                ò
```

```
The present invention describes genomic DNA which contains a BRCA2 gene where the first 12 nucleotides beginning exon 5 are 5'-TCCTGTTCTT-3' as in sequence (I) (see AAX03249), where nucleotides numbers 5782-5790 are GTTTGTTT as in sequence (IV) (see AAX30255), and where the last 20 nucleotides encoding exon 15 are 5'-CTGGGTTTCTCATAAACG-3' as in sequence (III) (see AAX30251) and the first 20 nucleotides CC as in sequence (III) (see AAX30251) and the first 20 nucleotides (III) (see AAX30253). Products and methods from the present invention can be used for identifying mutations in the BRCA2 gene leading to compare the present invention can be used for detection and gene therapy for breast and ovarian cancer. They can be used for detection and gene therapy for breast or ovarian cancer. They can be used in methods for monitoring disease progression, or for detection and gene and protein replacement conformation of the presence or quantifying the amount of a tumour growth inhibitor following such therapy. The BRCA2 protein, polypeptides, their functional equivalents, antibodies, and PNs may also the useful in the study of the characteristics of BRCA2 proteins, such as structure and function of BRCA2 in oncogenesis or subcellular concerns thuman BRCA2 (cmil).
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  972 ksdislnidkipeknndymnkwag-----ligpisnhsfggsfrtasnkeiklsehnikk 1026
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          :| |:|||
|||----eesgsqfeftqfrkpsyllqkstfevpenqmtllkttseecrdadlhvimnapsig 1174
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     1175 qvdsskqfegtveikrkfagllkndcnksasgyltde-----nevgfrgfysahgtkl 1227
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               1228 nvstealgkavklfsdieniseets---aevhpislsskchd----svvsmf---kien 1277
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             96 QRLFGKTAGRFDVYFISPQGLKFRSKSSLANYLHKNGETSL-KPEDFDFTVLSKRGIKSR 154
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               155 YKDCSMAALTSHL--QNQSNNSNWNLRTRSKCKKDVFMPPSSSSELQESRGLSNFTSTHL 212
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           285 D-RTVCISDAGACGETLSVTSEENSLVKKKERSLSSGSNFCSEQKTSGIINKFCSAKDSE 343
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       / Match 4.7%; Score 145; DB 20; Length 3418; Local Similarity 20.9%; Pred. No. 0.0017; les 96; Conservative 67; Mismatches 158; Indels 138; Gaps
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     240 -----GIPIKK-----TKKGCRKSCSGFVQSDSKRESVCNKADAESEPVAQKSQL 284
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        344 HNEKYEDTFLESEE------IGTKVEVVERKEHLHTDILKRGSEMDNN--CSP 388
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 36 KEDVAMELERVGEDEEQMMIKRSSECNPLLQEPIASAQFGATAGTECRKSVPCGWERVVK 95
                                  develop products for determining susceptibility to, detection and treatment of breast or ovarian cancer
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              213 LLKEDEGVDDVNFRKVRK-------PKGKVTILK-------
                 New coding sequence haplotypes of the human BRCA2 gene -
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 :| :| | : | | : :| | | : :| | | : :| | | | srnshnlefdgsdsskndtv---cihkdetdllftdqhn 1363
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               389 TRK----DFTGEKIFQEDTIPRTQIERRKTSLYFSSKYN 423
                                                                                                                                Claim 17; Page 68–75; 226pp; English.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             3418 AA;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             sednence
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             Query Match
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              Matches
ζ
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 ΟŻ
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     qq
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              Ö
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        q
```

Search completed: January 15, 2002, 13:50:15

Job time: 304 sec

```
January 15, 2002, 13:45:11; Search time 120.01 Seconds (without alignments) 511.063 Million cell updates/sec
                                                                                                                                                                           US-09-724-296-2
4349
1 MTKLPILGYWKIKGLVQPTR......KRLTARKRRSRKEEVEEDEK 828
                                                                                                                                                                                                                                                                                                                                                   522463
GenCore version 4.5
Copyright (c) 1993 - 2000 Compugen Ltd.
                                                                                                                                                                                                                                                                                                                                     Total number of hits satisfying chosen parameters:
                                                                                                                                                                                                                                                                                                     522463 seqs, 74073290 residues
                                                                OM protein - protein search, using sw model
                                                                                                                                                                                                                                                                                                                                                                                                                         Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries
                                                                                                                                                                                                                                                   BLOSUM62
Gapop 10.0 , Gapext 0.5
                                                                                                                                                                                                                                                                                                                                                                        Minimum DB seq length: 0 Maximum DB seq length: 20000000000
                                                                                                                                                                         Title:
Perfect score:
Sequence:
                                                                                                                                                                                                                                                Scoring table:
                                                                                                                                                                                                                                                                                                     Searched:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     Database
                                                                                                    Run on:
```

/gcgdata/geneseq/geneseqp/AA1981.DAT:*/gcgdata/geneseq/geneseqp/AA1982.DAT:*/gcgdata/geneseq/geneseqp/AA1983.DAT:*	gcgdata/geneseq/geneseqp/AA1984.DAT:* gcgdata/geneseq/geneseqp/AA1985.DAT:*	gogdata/geneseq/geneseqp/AA1986.DAT:* /gogdata/geneseq/geneseqp/AA1987.DAT:*	٠ د			SIDS8/9c9data/geneseq/geneseqp/AA1994.DAT:* /SIDS8/9c9data/geneseq/geneseqp/AA1995.DAT:*		gcgdata/geneseq/geneseqp/AA1997.DAT:v		: -	
			*	* * .	* *	* *	*	*	* *	. *	*
* * *	* *	* * -	* ··	* * .	* *	* *	*	* 1	* *	*	*
,			**************************************					eqp/Aa1981.DaT1: eqp/Aa1982.DaT: eqp/Aa1984.DaT: eqp/Aa1984.DaT: eqp/Aa1985.DaT: eqp/Aa1986.DaT: eqp/Aa1989.DaT: eqp/Aa1989.DaT: seqp/Aa1990.DaT: seqp/Aa1991.DaT: seqp/Aa1991.DaT: seqp/Aa1991.DaT: seqp/Aa1993.DaT: seqp/Aa1993.DaT: seqp/Aa1993.DaT: seqp/Aa1994.DaT: seqp/Aa1994.DaT:		groddata/geneseqp/Aal1982_Dat1:# groddata/geneseqp/Aal1982_Dat1:# groddata/geneseqp/Aal1981_Dat1:# groddata/geneseqp/Aal1981_Dat1:# groddata/geneseqp/aal1981_Dat1:# groddata/geneseqp/Aal1981_Dat1:# groddata/geneseqp/Aal1981_Dat1:# groddata/geneseqp/Aal1981_Dat1:# groddata/geneseqp/Aal1981_Dat1:# groddata/geneseqp/Aal1981_Dat1:# groddata/geneseqp/Aal1981_Dat1:# groddata/geneseqp/Aal1991_Dat1:# groddata/geneseqp/geneseqp/Aal1991_Dat1:#	eqp/Aa1981.DaT1: eqp/Aa1982.DaT1: eqp/Aa1984.DaT1: eqp/Aa1984.DaT1: eqp/Aa1984.DaT1: eqp/Aa1986.DaT1: eqp/Aa1989.DaT1: eqp/Aa1989.DaT1: seqp/Aa1999.DaT1: seqp/Aa1991.DaT1: seqp/Aa1991.DaT1: seqp/Aa1991.DaT1: seqp/Aa1991.DaT1: seqp/Aa1992.DaT1: seqp/Aa1992.DaT1: seqp/Aa1992.DaT1: seqp/Aa1993.DaT1: seqp/Aa1993.DaT1: seqp/Aa1995.DaT1: seqp/Aa1995.DaT1: seqp/Aa1995.DaT1: seqp/Aa1995.DaT1: seqp/Aa1995.DaT1: seqp/Aa1995.DaT1: seqp/Aa1995.DaT1: seqp/Aa1995.DaT1:

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

Description	GST signal peptide GST signal peptide GST signal peptide S. pombe delta228- Trypanosoma cruzi Trypanosoma cruzi Glutathione-S-tran Trypanosoma cruzi Glutathione-S-tran GST-SEP fusion pro Glutathione-S-tran Trypanosoma cruzi
SUMMARIES	AAY44498 AAY4450 AAY4499 AAR84565 AAR84568 AAY4501 AAR84569 AAY31954 AAR81731 AAR81731
D.B	21 21 21 16 16 21 16 20 20 16
% Query Match Length DB 1	828 600 371 564 643 229 472 379 1140 310
% Query Match	29.0 29.0 29.0 28.4 28.1 28.1 28.0 28.0 28.0
Score	4349 3092 1259.5 1259.5 1221.5 1221.5 1220 1217.5 1216.5
Result No.	110 110 110

08-JUN-1999; 99WO-US12910.

Human HCP/GST fusi Expression vector G17 antigen. Chim Turkey prolactin/G GST-Cre recombinas Amino acid sequenc	-B fus thione ntigen I tran thione acid I tran	transport transport transport transport transport transport transport transport	GST/G1 transport r
20000-		AAW74102 AAW74113 AAW74113 AAW74121 AAW74100 AAW74100 AAW74102 AAW74102	
21 16 16 22 22	119 119 120 120 20 20	200000000000000000000000000000000000000	000000000000000000000000000000000000000
824 856 362 426 324 324	27.1 35.2 28.2 23.2 24.7 24.7	247 248 248 248 248 248	2557 2557 2557 2557 2558 2558 2559
27.9 27.8 27.8 7.75 7.75	27.7 27.6 27.6 27.6 27.6 27.6 27.6	27.6 27.6 27.6 27.6 27.6 27.6 27.6	27.06 27.06 27.06 27.06 27.06 27.06 27.06
1212 1208.5 1207.5 1206 1205.5 1204.5	1203 1203 1201.5 1199.5 1199.5 1199.5	1999. 1999. 1999. 1999. 1999.	1199.5 1199.5 1199.5 1199.5 1199.5 1199.5 1199.5 1199.5
12 13 14 15 16 17	19 20 22 22 23 24 25	26 27 28 33 33 34 34	38 88 88 89 89 89 89 89 89 89 89 89 89 89

ALIGNMENTS

RESULT 1 AAY44498	
ID AAY44498 st	AAY44498 standard; Protein; 828 AA.
AAY44498;	
27-MAR-200¢	27-MAR-2000 (first entry) JAN UEVIDUS (UNS)
GST signal	GST signal peptide and S. pombe UVDE fusion protein. $\mathcal{L}U$
GST signal UV irradiat skin cancer	GST signal peptide; ultraviolet damage endonuclease; UVDE; Uvelp; UV irradiation; DNA damage; UV radiation damage; fusion protein; skin cancer; glutathione-S-transferase.
Schizosacch Synthetic.	Schizosaccharomyces pombe. Synthetic.
Key Peptide	Location/Qualifiers
Protein	/label= GST_signal_peptide 230828
Misc-difference	
/nc Misc-difference 85	/note= "Encoded by aaaa" ence 85
Misc-differ	/note= "Encoded by gtt"
	/note= "Encoded by gtt"
WO9963828-A1.	
16-DEC-1999.	
08-JUN-1999;	99WO-USI2910.

```
The present sequence is a fusion protein comprising the GST signal peptide and full length ultraviolet damage endonuclease (UVDE) from Schizosaccharomyces pombe. UVDE genes are isolated by introducing a foreign cDNA library into a repair-deficient E. coli strain and selecting for complemented cells by UV irradiation of the transformants. This provides stable endonuclease fragments for cleaving a double-stranded DNA molecule that has a distorted structure resulting from UV radiation damage, a photoproduct, an abasic site, mismatched nucleotide pairing, aplatinum diaduct, an intercalated molecule or alkylation of a nucleotide. Uvelp can be used in compositions for internal or topical application and as a therapeutic agent for skin cancers.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           481 YACLNTILRSMKERVFCSRTCRITTIQRDGLESVKQLGTQNVLDLIKLVEWNHNFGIHFM 540
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 541 RVSSDLFPFASHAKYGYTLEFAQSHLEEVGKLANKYNHRLTMHPGQYTQIASPREVVVDS 600
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       SKRIVFTILKQKAFKGNHPCVPSVCTITYSRFHCLPDTLKSLLPMSSKTTLSMLPQVNIG 300
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               301 ANSFSABTPVDLKKENETELANISGPHKKSTSTSTRKRARSSKKKATDSVSDKIDESVAS 360
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   361 YDSSTHLRRSSRSKKPVNYNSSSESEEGISKATKKVKQKEEEEYVEEVDEKSLKNESS 420
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  361 ydssthlrrssrskkpvnynssseseegiskatkkvkqkeeeeyveevdekslkness 420
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               480
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      480
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 KVDFLSKLPEMLKMFEDRLCHKTYLNGDHVTHPDFMLYDALDVVLYMDPMCLDAFPKLVC 180
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         FKKRIEAIPQIDKYLKSSKYIAWPLQGWQATFGGGDHPPKSDHLVPRGSMLRLLKRNIQI 240
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   61 IDGDVKLTQSMAIIRYIADKHNMLGGCPKERAEISMLEGAVLDIRYGVSRIAYSKDFETL 120
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       new truncated ultraviolet damage endonuclease for treatment of skin
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        1 MTKLPILGYWKIKGLVQPTRLLLEYLEEKYEEHLYERDEGDKWRNKKFELGLEFPNLPYY 60
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                SDEFEPVVPEQLETPISKRRRSRSSAKNLEKESTMNLDDHAPREMFDCLDKPIPWRGRLG
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     DB 21; Length 828;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     0; Indels
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   0; Mismatches
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              100.0%; Score 4349;
100.0%; Pred. No. 0;
                                                                                                                                                                                                                                                               Claim 7; Page 51; 133pp; English.
                                                                                                     Avery AM;
                     99US-0134752.
98US-0088521
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    Best Local Similarity 100.0
Matches 828; Conservative
                                                                                                   Kaur B,
                                                                                                                                            WPI: 2000-116417/10.
                                                          (UYEM-) UNIV EMORY.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        828 AA;
                                                                                                                                                                  N-PSDB; AAZ29858
08-JUN-1998;
                     18-MAY-1999;
                                                                                                     Doetsch PW,
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          Sequence
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     Query Match
                                                                                                                                                                                                                              cancers
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   181
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            241
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        421
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                121
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              181
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   241
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 301
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       61
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        δy
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          Óχ
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        ŏ
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 g
```

δ

g δλ g ò

g ò

ò

ó

```
The present sequence is a fusion protein comprising the GST signal peptide and delta228 truncated ultraviolet damage endomuclease (UVDE) from S. pombe. Delta228-UVDE is a 288 amino acid deletion of the N-terminal of the S. pombe uvel+ gene product. This is expressed in frame with a GST leader sequence. Stable endonuclease fragments can be produced
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             A new truncated ultraviolet damage endonuclease for treatment of skin
                                                                                                                                      721 GITQKQHYSESADPTAISGMKRRAHSDRVFDFPPCDPTMDLMIEAKEKEQAVFELCRRYE 780
                                                                                                                                                    601 AIRDLAYHDEILSRMKLNEQLNKDAVLIIHLGGTFEGKKETLDRFRKNYQRLSDSVKARL 660
                                         661 VLENDDVSWSVQDLLPLCQELNIPLVLDWHHHNIVPGTLREGSLDLMPLIPTIRETWTRK 720
                                                                                               541 rvssdlfpfashakygytlefaqshleevgklankynhrltmhpgqytqiasprevvvds 600
                                                                                                                                                                                                                                                                                                                                                                                                                                     UV radiation damage; photoproduct; abasic site; aplatinum diaduct; mismatched nucleotide pairing; nucleotide alkylation;
                                                                                                                                                                                                                                                                                                                                                                                                 GST signal peptide; glutathione-S-transferase signal peptide; ultraviolet damage endonuclease; UVDE; delta228-UVDE; repair-deficient E. coli strain; UV irradiation; DNA damage;
                                                                                                                                                                                                                                                                                                                                                                     GST signal peptide and delta228 S. pombe UVDE fusion protein.
                                                                                                                                                                                                        781 LONPPCPLEIMGPEYDQTRDGYYPPGAEKRLTARKRRSRKEEVEEDEK 828
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  /note= "Encoded by aaaa"
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           /note= "Encoded by gtt"
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      /note= "Encoded by gtt"
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           Disclosure; Page 56; 133pp; English.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       Location/Qualifiers
                                                                                                                                                                                                                                                                                            AAY44500 standard; Protein; 600 AA
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  Avery AM;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                         fusion protein; skin cancer.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    99WO-US12910.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 98US-0088521.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            99US-0134752.
                                                                                                                                                                                                                                                                                                                                                 (first entry)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    Schizosaccharomyces pombe
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   Kaur B,
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             WPI; 2000-116417/10.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          Misc-difference 147
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         UYEM-) UNIV EMORY.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  Misc-difference 85
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     Misc-difference 11
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          N-PSDB; AAZ29860
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   WO9963828-A1
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       08-JUN-1999;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 08-JUN-1998;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 L8-MAY-1999;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  Doetsch PW,
                                                                                                                                                                                                                                                                                                                                                  27-MAR-2000
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             16-DEC-1999
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    Synthetic.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     cancers -
                                                                                                                                                                                                                                                                                                                        AAY44500;
                                                                                                                                                                                                                                                                               AAY44500
                                                                                                                                                                                                                                                                                                                        pp
                                                                                                                                                                                                    Óλ
                                     δλ
                                                              QQ
                                                                                         οy
                                                                                                                    qq
                                                                                                                                              ŏ
```

S. pombe delta228-UV damage endonucleas.

(first entry)

27-MAR-2000

AA

AAY44499 standard; Protein; 371

for cleaving a double-stranded DNA molecule that has a distorted structure resulting from UV radiation damage, a photoproduct, an abasic site, mismatched nucleotide pairing, aplatinum diaduct, an intercalated molecule or alkylation of a nucleotide. Uvelp can be used in compositions for internal or topical application and as a therapeutic agent for skin IDGDVKLTQSMAIIRYIADKHNMLGGCPKERAEISMLEGAVLDIRYGVSRIAYSKDFETL 120 KVDFLSKLPEMLKMFEDRLCHKTYLNGDHVTHPDFMLYDALDVVLYMDPMCLDAFPKLVC 180 SKRIVFTILKQKAFKGNHPCVPSVCTITYSRFHCLPDTLKSLLPMSSKTTLSMLPQVNIG 300 Indels 228; Gaps 1 MTKLPILGYWKIKGLVQPTRLLLEYLEEKYEEHLYERDEGDKWRNKKFELGLEFPNLPYY FKKRIEAIPQIDKYLKSSKYIAWPLQGWQATFGGGDHPPKSDHLVPRGSMLRLLKRNIQI Length 600; Score 3092; DB 21; Pred. No. 5.6e-225; Mismatches ; 71.18; 72.58; Conservative Similarity 600 AA; Matches 600; Sequence Query Match cancers Best Local 121 181 181 230 241 8888888888qq g ò ò g q g

ò

ò

ò

420 372 301 ANSFSAETPVDLKKENETELANISGPHKKSTSTSTRKRARSSKKKATDSVSDKIDESVAS 360 540 600 432 VLENDDVSWSVQDLLPLCQELNIPLVLDWHHHNIVPGTLREGSLDLMPLIPTIRETWTRK 720 GITOKQHYSESADPTAISGMKRRAHSDRVFDFPPCDPTMDLMIEAKEKEQAVFELCRRYE 780 361 YDSSTHLRRSSRSKKPVNYNSSSESEEQISKATKKVKQKEEEETVEEVDEKSLKNESS 421 SDEFEPVVPEQLETPISKRRSRSSAKNLEKESTMNLDDHAPREMFDCLDKPIPWRGRLG -----ddhapremfdcldkpipwrgrlg YACLNTILRSMKERVFCSRTCRITTIQRDGLESVKQLGTQNVLDLIKLVEWNHNFGIHFM RVSSDLFPFASHAKYGYTLEFAQSHLEEVGKLANKYNHRLTMHPGQYTQIASPREVVVDS AIRDLAYHDEILSRMKLNEQLNKDAVLIIHLGGTFEGKKETLDRFRKNYQRLSDSVKARL 230 181 541 313 601 661 g ò g g ò qq Q g ò ò ò à g

.; 0

TLREGSLDLMPLIPTIRETWTRKGITQKQHYSESADPTAISGMKRRAHSDRVFDFPPCDP 757

869

638 KKETLDRFRKNYQRLSDSVKARLVLENDDVSWSVQDLLPLCQELNIPLVLDWHHHNIVPG 697

 $_{Q}$ qq δ

552

LQNPPCPLEIMGPEYDQTRDGYYPPGAEKRLTARKRRSRKEEVEEDEK 828

553

AAY44499

493 781

g ð

ò

The present sequence is delta228-UV damage endonuclease. Delta228-UVDE is a 288 amino acid deletion of the N-terminal of the S. pombe uvel+ gene product. This is expressed in frame with a GST leader sequence to generate a fusion protein. This provides s'able endonuclease fragments for cleaving a double-stranded DNA molecui. that has a distorted structure resulting from UV radiation damage, a photoproduct, an abasic site, mismatched nucleotide pairing, aplatinum diaduct, an intercalated molecule or alkylation of a nucleotide. Uvelp can be used in compositions for internal or topical application and as a therapeutic agent for skin new truncated ultraviolet damage endonuclease for treatment of skin 458 DDHAPREMFDCLDKPIPWRGRLGYACLNTILRSMKERVFCSRTCRITTIQRDGLESVKQL 517 518 GTQNVLDLIKLVEWNHNFGIHFMRVSSDLFPFASHAKYGYTLEFAQSHLEEVGKLANKYN 577 0; Gaps Delta228-UVDE; ultraviolet damage endonuclease; GST signal peptide; quitathione-S-transferase signal peptide; uvel+ gene product; UV irradiation; DNA damage; UV radiation damage; photoproduct; abasic site; aplatinum diaduct; mismatched nucleotide pairing; HRLTMHPGQYTQIASPREVVVDSAIRDLAYHDE1LSRMKLNEQLNKDAVLIIHLGGTFEG 45.5%; Score 1979; DB 21; Length 371; Indels 0 Pred. No. 3e-141; Mismatches nucleotide alkylation; skin cancer. Claim 13; Page 53; 133pp; English. 45.5%, 100.0%; Pre 0; Avery AM; 99WO-US12910 99US-0134752 Schizosaccharomyces pombe. Conservative Doetsch PW, Kaur B, WPI; 2000-116417/10 (UYEM-) UNIV EMORY. Best Local Similarity 371 AA; N-PSDB; AAZ29859 WO9963828-A1 08-JUN-1999; 18-MAY-1999; 08-JUN-1998; 16-DEC-1999 Matches 371; Sequence Query Match cancers cancers 61 578 δ pp ò qq δ qα

Tue Jan 15 14:12:04 2002

qq οy qq δ g

```
AAR84565-R84569 are polypeptides of the TCR27 protein of T.cruzi. The proteins are all fusion products with glutathione-S-transferase (GST) and some contain a linker sequence. The TCR27 protein comprises conserved 14 aa sequence and a 68 aa C.terminal region. This sequence encodes the GST sequence, the Ag15 polypeptide contq. 16 of the 69 repeat units and also contains a linker sequence. The TCR27 polypeptides of the invention are useful for the diagnosis of chagas disease (American Trypanosomiasis), they are capable of detecting anti-T.cruzi antibodies, or for blood screening. The TCR27 protein has epitopes to which most T.cruzi infected individuals have antibodies. The TCR27 polypeptides will not react with serum from
                                                         241 tlregsldlmpliptiretwtrkgitqkqhysesadptaisgmkrrahsdrvfdfppcdp 300
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              patients with leishmaniasis, schistosomiasis, or autoimmune disease and are hence less likely to cause false positives in diagnosis.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     New polypeptide(s) from the TCR27 protein of Trypanosoma cruzi - a immunoassay reagent for specific diagnosis of Chagas disease, also related nucleic acid and transformed cells
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    329..552
/label= repeat_region
/note= "16 of 69 repeat units of 14 amino acids"
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      TCR27; Chagas disease; repeat unit; diagnosis; blood screening; recombinant; fusion protein; glutathione-S-transferase.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             Trypanosoma cruzi TCR27 polypeptide, Ag15.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         Disclosure; Page 30-31; 68pp; English.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     Location/Qualifiers
                                                                                                                                                                                                                                                                                                                                                                                                                  AAR84565 standard; Protein; 564 AA.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              95WO-US03191.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        94US-0216894,
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   (first entry)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     (KIRC/) KIRCHHOFF L V. (OTSU/) OTSU K.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   Kirchhoff LV, Otsu K;
                                                                                                                                                                                          818 SRKEEVEEDEK 828
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                WPI; 1995-344618/44.
                                                                                                                                                                                                                               564 AA;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            Trypanosoma cruzi.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  N-PSDB; AAT05329.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    WO9525797-A1.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              20-MAR-1995;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        24-MAR-1994;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 28-SEP-1995.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                09-MAY-1996
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   Sednence
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   AAR84565;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          Region
                                                                                                                                                                                                                                                                                                                                                                                       AAR84565
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            Key
```

Score 1259.5; DB 16; Length 564; Pred. No. 1.2e-86;

29.0%; 57.1%;

Query Match Best Local Similarity

```
12;
                                                                                                                                                                                                                                                                                                          242 k-emtklrekvkkaekekldainratkleeernqaykaahkaeeekaktfgrlitfesen 300
                                                                                                                185 IEAIPQIDKYLKSSKYIAWPLQGWQATFGGGDHPPKSDHLVPRG--SMLRLLKRNIQISK 242
                                                                                                                                                                                                                                            243 RIVETILKOKAFKGNHPCVPSVCTIT-----YSREHCLPD----TLKSLLPMSSKT 289
                                                                                                                                                                                                                                                                                                                                                        290 -TLSMLPQVNIG----ANSFSAETPVDLKKENETELANISGPHKKSTSTSTRKRARSSK 343
                                                                                                                                                                                                                                                                                                                                                                         34 KKATDSV----SDKIDESVASYDSSTHLRRSSRSKKPVNYNSSSESESEEQISKATKKVK 399
                                                                                                                                                                                                                                                                                                                                                                                                                                                     359 rkaaeaakavgtekgraaeatkvaeaekgkaaeaakav-----etekgraaeatkva 410
                                                                                                                                                            125 LSKLPEMLKMFEDRLCHKTYLNGDHVTHPDFMLYDALDVVLYMDPMCLDAFPKLVCFKKR 184
                                                                                               65 VKLTQSMAIIRYIADKHNMLGGCPKERAEISMLEGAVLDIRYGVSRIAYSKDFETLKVDF 124
 45; Gaps
                                                 3 pilgywkikglvqptr1lleyleekyeehlyerdegdkwrnkkfelglefpnlpyyidgd 62
                                5 PILGYWKIKGLVQPTRLLLEYLEEKYEEHLYERDEGDKWRNKKFELGLEFPNLPYYIDGD 64
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    New polypeptide(s) from the TCR27 protein of Trypanosoma cruzi - a immunoassay reagent for specific diagnosis of Chayas disease, also
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         329..552
/label= repeat_region
/note= "16 of 69 repeat units of 14 amino acids"
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          TCR27; Chagas disease; repeat unit; diagnosis; blood screening; recombinant; fusion protein; glutathione-S-transferase.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          400 QKEEEEYVEEVDEKSLKNESSSDEFEPVVPEQLETPISKRRSRSSAKNLEKE 452
 55; Mismatches 103; Indels
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               Trypanosoma cruzi TCR27 polypeptide, Ag15.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            Location/Qualifiers
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 AAR84568 standard; Protein; 643 AA.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         94US-0216894.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         95WO-US03191.
   Matches 270; Conservative
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                (first entry)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          (KIRC/) KIRCHHOFF L V.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        Kirchhoff LV, Otsu K;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      WPI; 1995-344618/44.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               Trypanosoma cruzi.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      N-PSDB; AAT05332.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           (OTSU/) OTSU K.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            WO9525797-A1.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         20-MAR-1995;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         24-MAR-1994;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                09-MAY-1996
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          28-SEP-1995.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 AAR84568;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              Region
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   AAR84568
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              ò
                                                                                                                                                                                           QQ
                                                                                                                                                                                                                             ò
                                                                                                                                                                                                                                                        Q
                                                                                                                                                                                                                                                                                          QY
                                                                                                                                                                                                                                                                                                                         QΩ
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      Öλ
```

```
AAR84565-R84569 are polypeptides of the TCR27 protein of T.cruzi

The proteins are all fluion products with glutathione-S-transferase

(GST) and some contain a linker sequence. The TCR27 protein comprises

a 95 amino acid (aa) N-terminal region; 69 repeats of a highly

conserved 14 as sequence and a 68 as C-terminal region. This sequence

encodes the GST sequence, the A944 polypeptide contg. 16 of the 69

repeat units and also contains the amino and carboxy terminal

c peptides of TCR27. The TCR27 polypeptides of the invention are useful

for the diagnosis of Chagas disease (American Trypanosomiasis), they

are capable of detecting anti-T.cruzi antibodies; or for blood

c screening. The TCR27 protein has epitopes to which most T.cruzi

infected individuals have antibodies. The TCR27 polypeptides will not

react with serum from patients with leishmaniasis, schistosomiasis,

c or autoimmune disease and are hence less likely to cause false
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         242 k-emtklrekvkkaekekldainratkleeernqaykaahkaeeekaktfgrlitfesen 300
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             -TLSMLPQVNIG----ANSFSAETPVDLKKENETELANISGPHKKSTSTSTRKRARSSK 343
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       65 VKLTQSMAIIRYIADKHNMLGGCPKERAEISMLEGAVLDIRYGVSRIAYSKDFETLKVDF 124
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       LSKLPEMLKMFEDRLCHKTYLNGDHVTHPDFMLYDALDVVLYMDPMCLDAFPKLVCFKKR 184
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              IEAIPQIDKYLKSSKYIAWPLQGWQATFGGGDHPPKSDHLVPRG--SMLRLLKRNIQISK 242
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          ---YSRFHCLPD----TLKSLLPMSSKT 289
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        344 KKATDSV----SDKIDESVASYDSSTHLRRSSRSKKPVNYNSSSESESEEQISKATKKVK 399
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  45; Gaps
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               5 PILGYWKIKGLVQPTRLLLEYLEEKYEEHLYERDEGDKWRNKKFELGLEFPNLPYYIDGD 64
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  29.0%; Score 1259.5; DB 16; Length 643; 57.1%; Pred. No. 1.4e-86; ative 55; Mismatches 103; Indels 45;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         400 QKEEEEYVEEVDEKSLKNESSSDEFEPVVPEQLETPISKRRRSRSSAKNLEKE 452
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 GST signal peptide; Glutathione-S-transferase signal peptide; ultraviolet damage endonuclease; UVDE; delta228-UVDE; UV irradiation; DNA damage; UV radiation damage; photoproduct;
related nucleic acid and transformed cells
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               Glutathione-S-transferase signal peptide.
                                     Disclosure; Page 40-41; 68pp; English.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      243 RIVFTILKQKAFKGNHPCVPSVCTIT-----
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  AAY44501 standard; Protein; 229 AA.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                         Matches 270; Conservative
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         27-MAR-2000 (first entry)
                                                                                                                                                                                                                                                                                                                                            positives in diagnosis.
                                                                                                                                                                                                                                                                                                                                                                                                                                                           Similarity
                                                                                                                                                                                                                                                                                                                                                                                643 AA;
                                                                                                                                                                                                                                                                                                                                                                                  Sequence
                                                                                                                                                                                                                                                                                                                                                                                                                                     Query Match
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        AAY44501;
                                                                                                                                                                                                                                                                                                                                                                                                                                                           Sest Local
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              9
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         123
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                183
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     125
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               290
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              AAY44501
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              RESULT
ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       pp
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              q
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      ð
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         a
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 셤
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 οy
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        ò
```

```
ultraviolet damage endonuclease from S. pombe uvel+ gene product is expressed in frame with the GST leader sequence to generate a fusion protein. This provides stable endonuclease fragments for cleaving a double-stranded DNA molecule that has a distorted structure resulting from UV radiation damage, a photoproduct, an abasic site, mismatched nucleotide pairing, aplatinum diaduct, an intercalated molecule or alkylation of a nucleotide. Uvelp can be used in compositions for internal or topical application and as a therapeutic agent for skin
                                                                                                                                                                                                                                                                                                                                                                                                    A new truncated ultraviolet damage endonuclease for treatment of skin
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            61 IDGDVKLTQSMAITRYIADKHNMLGGCPKERAEISMLEGAVLDIRYGVSRIAYSKDFETL 120
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             121 KVDFLSKLPEMLKMFEDRLCHKTYLNGDHVTHPDFMLYDALDVVLYMDPMCLDAFPKLVC 180
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 Gaps
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        1 MTKLPILGYWKIKGLVQPTRLLLEYLEEKYEEHLYERDEGDKWRNKKFELGLEFPNLPYY 60
                                                                                                                                                                                                                                                                                                                                                                                                                                                                          The present sequence is GST signal peptide. Delta228, a truncated
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  0;
 abasic site; aplatinum diaduct; mismatched nucleotide pairing;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    Length 229;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         181 FKKRIEAIPQIDKYLKSSKYIAWPLQGWQATFGGGDHPPKSDHLVPRGS 229
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        Indels
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    28.4%; Score 1237; DB 21;
100.0%; Pred. No. 1.6e-85;
tive 0; Mismatches 0;
                                                                                            'note= "Encoded by aaaa"
                                                                                                                                                   /note= "Encoded by gtt"
                                                                     Location/Qualifiers
                                                                                                                          ρλ
              nucleotide alkylation; skin cancer.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     AAR84569 standard; Protein; 472 AA.
                                                                                                                                                                                                                                                                                                                                                                                                                                              Claim 9; Page 58; 133pp; English.
                                                                                                                         /note= "Encoded
                                                                                                                                                                                                                                                                                                                                     Avery AM;
                                                                                                                                                                                                                                     99WO-US12910
                                                                                                                                                                                                                                                              98US-0088521
                                                                                                                                                                                                                                                                            990S-0134752
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              Conservative
                                                                                                                                                                                                                                                                                                                                  Doetsch PW, Kaur B,
                                                                                                                                                                                                                                                                                                                                                            WPI; 2000-116417/10.
                                                                                                                                   Misc-difference 147
                                                                                                                                                                                                                                                                                                      (UYEM-) UNIV EMORY.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 Best Local Similarity
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            229 AA;
                                                                               Misc-difference
                                                                                                         Misc-difference
                                                                                                                                                                                                                                                                                                                                                                            N-PSDB; AAZ29861
                                                                                                                                                                              WO9963828-A1
                                                                                                                                                                                                                                  08-JUN-1999;
                                                                                                                                                                                                                                                              08-JUN-1998;
                                                                                                                                                                                                                                                                              18-MAY-1999;
                                                                                                                                                                                                       16-DEC-1999
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             Matches 229;
                                        Synthetic.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            Seguence
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    Query Match
                                                                                                                                                                                                                                                                                                                                                                                                                    cancers
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  cancers
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         AAR84569
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           RESULT
ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  Ω
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      QQ
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   õ
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             q
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      X ID
                                                                                                                                                                                                                                                                                                                                             12;
```

ö

```
AAY31954;
                                                                                                                                                                                                                                                                                                                                                                                                                     AAY31954
                                                                                                                                                                                                                                                                                                                                                                                                  RESILLT
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                qq
                                                                                                                                                                                                                                                                                                οy
                                                                                                                                    δŏ
                                                                                                                                                                                                               QΛ
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            9:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            AAR84565-R84569 are polypeptides of the TCR27 protein of T.cruzi
C The proteins are all fusion products with glutathione-S-transferase
(GST) and some contain a linker sequence. The TCR27 protein comprises
a 95 amino acid (aa) N-terminal region; 69 repeats of a highly
c onserved 14 as sequence and a 68 as C-terminal region. This sequence
c conserved 14 as sequence, the Ag8 polypeptide contg. 15 of the 69
c repeat units and also contains a linker sequence. The TCR27
polypeptides of the invention are useful for the diagnosis of Chagas
C disease (American Trypanosomiasis), they are capable of detecting
anti-T.cruzi antibodies; or for blood screening. The TCR27 protein
has epitopes to which most T cruzi infected individuals have
c antibodies. The TCR27 polypeptides will not react with serum from
patients with leishmaniasis, schistosomiasis, or autoimmune disease
and are hence less likely to cause false positives in diagnosis.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        63 vkltqsmaiiryiadkhnmlggcpkeraeismlegavldirygvsriayskdfetlkvdf 122
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        185 IEAIPQIDKYLKSSKYIAWPLQGWQATFGGGDHPPKSDHLVPRGSMLRLLKRNIQISKRI 244
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              125 LSKLPEMLKMFEDRLCHKTYLNGDHVTHPDFMLYDALDVVLYMDPMCLDAFPKLVCFKKR 184
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 65 VKLTQSMAIIRYIADKHNMLGGCPKERAEISMLEGAVLDIRYGVSRIAYSKDFETLKVDF 124
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          3 pilgywkikglvqptrllleyleekyeehlyerdegdkwrnkkfelglefpnlpyyidgd 62
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    5 PILGYWKIKGLVQPTRLLLEYLEEKYEEHLYERDEGDKWRNKKFELGLEFPNLPYYIDGD 64
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             New polypeptide(s) from the TCR27 protein of Trypanosoma cruzi - a immunoassay reagent for specific diagnosis of Chagas disease, also related nucleic acid and transformed cells
                                                                                                                                                                                                                                        241..450
/label= repeat_region
/note= "15 of 69 repeat units of 14 amino acids"
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   28.1%; Score 1221.5; DB 16; Length 472; 56.7%; Pred. No. 6.6e-84; tive 54; Mismatches 110; Indels 31;
                                                                                                               TCR27; Chagas disease; repeat unit; diagnosis; blood screening; recombinant; fusion protein; glutathione-S-transferase.
                                                                          Trypanosoma cruzi TCR27 polypeptide, Ag8.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            Disclosure; Page 44-45; 68pp; English.
                                                                                                                                                                                                                     Location/Qualifiers
                                                                                                                                                                                                                                                                                                                                                                                                  95WO-US03191.
                                                                                                                                                                                                                                                                                                                                                                                                                                           94US-0216894.
                                      09-MAY-1996 (first entry)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            Matches 255; Conservative
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   (KIRC/) KIRCHHOFF L V.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               Kirchhoff LV, Otsu K;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   WPI; 1995-344618/44.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           Best Local Similarity
                                                                                                                                                                               Trypanosoma cruzi.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         N-PSDB; AAT05333
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           (OTSU/) OTSU K.
                                                                                                                                                                                                                                                                                                                                                                                                                                             24-MAR-1994;
                                                                                                                                                                                                                                                                                                                       WO9525797-A1
                                                                                                                                                                                                                                                                                                                                                                                                      20-MAR-1995;
                                                                                                                                                                                                                                                                                                                                                              28-SEP-1995.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                Sequence
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       Query Match
AAR84569;
                                                                                                                                                                                                                                             Region
```

ò g δλ g Db

```
305 SAETP--VDLKKENETELANISGPHKKSTSTRTKRARSSKKKATDSVSDKIDESVASYD 362
                                                                                                                                                                                                                                                                                                                            363 SSTHLRRSSRSKKPVNYNSSSESESEQISKATKKVKQKEEEEYVEEVDEKSLKNESSSD 422
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  342 kaaeatkvaeaekqkaaeatkvaeaekgkaaeatkvaeaekqkaaeatkvaeaekgkaa- 400
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   native Weel. The fusion protein was expressed in Escherichia coli cells that had been transformed with a pGEX vector into which a PGE fragment encoding amino acids 11-152 of Weel had been cloned. Weel participates in the regulation of cell cycle checkpoints during mitosis by catalysing phosphorylation of the cyclin-dependent kinese that initiates mitosis, Cdc2, Weel is itself phosphorylated
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       by cdc1. The fusion protein is useful in assays for evaluating Weel function and for screening for modulators of Weel and Cds1. It can be used in combination with other components of the cell cycle, particularly Cds1, Chk1, Cdc2, Cdc25 and other checkpoint proteins, in methods to identify and characterise regulatory pathways and to identify compounds which modulate components in the pathways,
                                                                                                                      245 VFTILKQKAFKGNHPCVPSVCTITYSRFHCLPDTLKSLLPMSSKTTLSMLPQVNIGANSF 304
                                                                                                                                                                                                  238 a-eaekgkaaeat-----kvaeaekgraaeatk--vaeaekgkaaeatkvaeaekgr
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     The present sequence represents a fusion protein, termed GST:Weel(11-152), comprising glutathione-S-transferase and a truncated Weel protein composed of amino acid residues 11-152 of
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   Truncated fusion proteins and related polynucleotides useful for screening for compounds which modulate Weel function \,
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    GST:Weel(11-152); glutathione-S-transferase; GST; WEE1; human; mitosis; cell cycle checkpoint; cyclin-dependent kinase; Cds1.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             Glutathione-S-transferase:Weel(11-152) fusion proein.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          423 EFEPVVPEQLETPISKRRRSRSSAKNLEKE 452
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  401 -----eatkvaeaekqkaaeatkvaeae 423
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              Claim 6; Page 40-42; 46pp; English.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    AAY31954 standard; Protein; 379 AA
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         Russell P, Boddy MN, Funari B;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       Chimeric - Schistosoma mansoni.
Chimeric - Homo sapiens.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 99WO-US06704.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             98US-0079752.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         21-DEC-1999 (first entry)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         (SCRI ) SCRIPPS RES INST.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          WPI; 1999-591103/50.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    particularly Cdcl.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    N-PSDB; AAZ20067
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         WO9949061-A1.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             26-MAR-1999;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 27-MAR-1998;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     30-SEP-1999.
```

```
glutathione S trinferase (GST) and stroling at the linker sequence between the GST and Stroling at the linker sequence between the GST and Stroling at the linker sequence was isolated from Molt 4 human T-cell leukaemia cells (ATCC CRL 1582). The sep gene was amplified in five fragments using the primers given in AAT00756 69. The amplified SEP gene was inserted into a vector already protein was used in the isolation protein expressed. The fusion protein was used in the isolation of a protein of mammalian origin which blunds a glutathione-S-transferase (GST)-FK506 binding protein (FKBP) rapamycin complex. The FKBP-rapamycin binding protein and corresponding DNA was isolated from Molt4 cells using a complex of the fusion protein of a protein and corresponding DNA was isolated from Molt4 cells using a complex of the fusion protein GST-FKBPI and rapamycin. The isolated proteins have
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 65 VKLTQSMAIIRYIADKHNMLGGCPKERAEISMLEGAVLDIRYGVSRIAYSKDFETLKVDF 124
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  404 ekkklrhasga----nitnattaattatttastegsnseseaestensp-tpspl 456
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      125 LSKLPEMLKMFEDRLCHKTYLNGDHVTHPDFMLYDALDVVLYMDPMCLDAFPKLVCFKKR 184
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    185 IEAIPQIDKYLKSSKYIAWPLQGWQATFGGGDHPPKSDHLVPRGSMLRLLKRNIQISKRI 244
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    245 VFTILKQK-AFKGNHPCVPSVCTITYSRFHCLPDTLKSLLPMSSKTTL----SMLPQVN 298
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     289 yaymkumwksarkidafqhmqhfvqtmqqqaqhaiatedqqhkqelhklmarcflklgew 348
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             329 ----KSTSTRKRARSSKKKATDSVSDKIDES-----VASYDSSTHLRRSSR--- 372
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                349 glnlgginestipkvlgyysaateh----drswykawhawavmnfeavlhykhqnqprd 403
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     373 SKKPVNYNSSSESESEEQISKATKKVKQKEEEEPYVEEVDEKSLKNESSSDEFEPVVPEQL 432
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              -REMFDCLDKPIPWRGRLGYACLNTILRSMKERVFCSRTC---RITT 505
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                506 IQRDGLESVKQLGTQNVLDLIKLVEWNHNFGIHFMRVSSDLFPFAS--HAKYGYTLEFAQ 563
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 577 arhnaanki-----lknmcehsntlvqqammvseelirvailwh-----emwh 619
                                                                                                                                                                                                                                                                                                      fusion protein GST-FKBPI2 and rapamycin. The isolated proteins have molecular weights of 125, 148, 208 and 210 kD. They can be used for identifying an immunomodulatory, or an antitumour agent. They can a be used in the detection of rapamycin, rapamycin analogues or metabolites when complexed with FKBP. Antisense DNA can be used to modulate the immune system of a mammal.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     Matches 344; Conservative 103; Mismatches 261; Indels 241; Gaps
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            5 PILGYWKIKGLVQPTRLLLEYLEEKYEEHLYERDEGDKWRNKKFELGLEFPNLPYYIDGD
   glutathione-S-transferase-FK506 binding protein-rapamycin complex
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          517 gvkaigidtwlgvipgliaridtprplvgrlihglltdigryhpgaliypltvasksttt
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   DB 16; Length 1140;
                                                                                This sequence represents a fragment of a fusion protein between
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    299 IG--ANSFSAETPVD------LKKENETELANISGPHK----
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             28.0%; Score 1217.5; DB 36.2%; Pred. No. 4.8e-83;
                                           Example 2; Page 29-35; 44pp; English.
                                                                                                                                                                                                                                                                                                                                                                                                                                            1140 AA;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               Query Match
Best Local Similarity
                                                                                                                                                                                                                                                                                                                                                                                                                                            Seguence
       QQ
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     δ
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             qq
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           õ
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      QΩ
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            Q
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     Db
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    õ
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           ŏ
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                qq
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   ŏ
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         qq
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              \delta
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               q
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       ó
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          a
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                ó
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 QQ
                                                                                                ω
..
                                                                                                                                                                                                    65 VKLTQSMAIIRYIADKHNMLGGCPKERAEISMLEGAVLDIRYGVSRIAYSKDFETLKVDF 124
                                                                                                                                                                                                                        242 glrrsgrsmnlnra---tll-----tll-----------apptpsslydannst 273
                                                                                                                                                                                                                                                                           125 LSKLPEMLKMFEDRLCHKTYLNGDHVTHPDFMLYDALDVVLYMDPMCLDAFPKLVCFKKR 184
                                                                                                                                                                                                                                                                                                                                                                                                                                231 -LRLLKRNIQISKRIVFTILKQKAFKGNHPCVPSVCTITYSRFHCLPDTLKSLLPMSSKT 289
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         290 TLSMLPQVNIGANSFSAETPVDLKKENETELANISGPHKKSTS------332
                                                                                                Gaps
                                                                                                                           5 PILGYWKIKGLVQPTRLLLEYLEEKYEEHLYERDEGDKWRNKKFELGLEFPNLPYYIDGD 64
                                                                                                                                                 274 -----sstssqkp-----ntsftslfgprkqttsspsfshaaplhplspps 314
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                333 -TSTRKRARSSKKKATDSVSDK---IDESVASYDSSTHLRRSSRSKKPVNYNSSSESE 386
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 315 fthsqpqiqaqpvprrpslfdrpnlvsrsssrlgdspslspvaqvanpihhtapspsd 372
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               Glutathione-S-transferase, GST; FK506 binding protein; FKBP; PCR; rapamycin; FKBP-rapamycin binding protein; Molt4 cell; amplify; fusion protein; GST-FKBP12; immunomodulatory agent; primer; antitumour agent; detection; antisense DNA; immune system.
                                                         Length 379;
                                                                                              Indels
                                                                                                                                                                                                                                                                                                                                                      IEAIPQIDKYLKSSKYIAWPLQGWQATFGGGDHPPKSDHLVPRGSM----
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             New effector proteins of rapamycin - which bind to a
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   Molnar-Kimber KL;
                                                          DB 20;
                                                   28.1%; Score 1220; DB 20
60.3%; Pred. No. 6.2e-84;
Live 32; Mismatches 50
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                AAR81731 standard; Protein; 1140 AA.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 Chen Y, Failli AA,
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              GST-SEP fusion protein fragment.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      (AMHP ) AMERICAN HOME PROD CORP. (UYCO ) UNIV COLUMBIA NEW YORK.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             95US-0384524.
94US-0207975.
94US-0312023.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            95EP-0301475.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        29-MAY-1996 (first entry)
                                                                                            Conservative
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      1995-346091/45
                                                                        Local Similarity
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        N-PSDB; AAT00771
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          07-MAR-1995;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             Caggiano TJ,
Nakanishi K;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              Homo sapiens
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             13-FEB-1995;
08-MAR-1994;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  26-SEP-1994;
                                                                                          Matches 252;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 EP676471-A2
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    11-0CT-1995
Sequence
                                                       Query Match
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       AAR81731;
                                                                                                                                                                                                                                                                                                                                                        185
                                                                                                                                                               g
                                                                                                                                                                                                      ò
                                                                                                                                                                                                                                        q
                                                                                                                                                                                                                                                                             ò
                                                                                                                                                                                                                                                                                                                g
                                                                                                                                                                                                                                                                                                                                                                                         qq
                                                                                                                                                                                                                                                                                                                                                                                                                                                                  g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           Dp
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  ò
                                                                                                                               ô
                                                                                                                                                                                                                                                                                                                                                      ò
                                                                                                                                                                                                                                                                                                                                                                                                                                ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       a
```

34;

They can also

particularly Cdcl.

9

```
The present sequence represents a fusion protein, termed GST:Weel(11-72), comprising glutathione-S-transferase and a curnoated Weel protein composed of amino acid residues 11-72 of native Weel. The fusion protein was expressed in Escherichia calis cells that had been transformed with a pGEX vector into which can per fragment encoding amino acids 11-72 of Weel had been cloned. Weel participates in the regulation of cell cycle checkpoints during mitosis by catalysing phosphorylation of the cyclin-dependent mitosis by catalysing phosphorylation of the cyclin-dependent through that includes mitosis, CdC2. Weel is useful in assays for evaluating Weel function and for screening for modulators of Weel and Cds1. It can be used in combination with other components of the call cycle, particularly Cds1, Chk1, Cdc2, Cdc25 and other checkpoint proteins, in methods to identify and characterise regulatory pathways and to identify compounds which modulate components in the pathways.
                                                  617 LNEQLNKDAVLIIHLGGTFEGKKETLDRFRKNYQRLSDSVKARLVLENDDVSWSVQDLLP 676
                                                                                                                                                                                                                                       675 wcrkymk----sgnvkdltgawdlyyhvfrriskglpgltslelgyvs----pkll 722
                                                                                                                                                                                                                                                                                                                       677 LCOELNIPLVLDWHHHNIVPGTL--REGSLDLMPLIPTIRETWTRKGITOKQHYSE---- 730
                                                                                                                                                                                                                                                                                                                                                                                                  723 mcrdlel-----avpgtydpngpiirigsiapslgv----itskgrprkltlm 766
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  ------HSD 747
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                : | | |: :: | | :: | | :: | | :: | | :: | | :: | | :: | | :: | :: | :: | :: | :: | :: | :: | :: | :: | :: | :: | :: | :: | :: | :: | :: | :: | :: | :: | :: | :: | :: | :: | :: | :: | :: | :: | :: | :: | :: | :: | :: | :: | :: | :: | :: | :: | :: | :: | :: | :: | :: | :: | :: | :: | :: | :: | :: | :: | :: | :: | :: | :: | :: | :: | :: | :: | :: | :: | :: | :: | :: | :: | :: | :: | :: | :: | :: | :: | :: | :: | :: | :: | :: | :: | :: | :: | :: | :: | :: | :: | :: | :: | :: | :: | :: | :: | :: | :: | :: | :: | :: | :: | :: | :: | :: | :: | :: | :: | :: | :: | :: | :: | :: | :: | :: | :: | :: | :: | :: | :: | :: | :: | :: | :: | :: | :: | :: | :: | :: | :: | :: | :: | :: | :: | :: | :: | :: | :: | :: | :: | :: | :: | :: | :: | :: | :: | :: | :: | :: | :: | :: | :: | :: | :: | :: | :: | :: | :: | :: | :: | :: | :: | :: | :: | :: | :: | :: | :: | :: | :: | :: | :: | :: | :: | :: | :: | :: | :: | :: | :: | :: | :: | :: | :: | :: | :: | :: | :: | :: | :: | :: | :: | :: | :: | :: | :: | :: | :: | :: | :: | :: | :: | :: | :: | :: | :: | :: | :: | :: | :: | :: | :: | :: | :: | :: | :: | :: | :: | :: | :: | :: | :: | :: | :: | :: | :: | :: | :: | :: | :: | :: | :: | :: | :: | :: | :: | :: | :: | :: | :: | :: | :: | :: | :: | :: | :: | :: | :: | :: | :: | :: | :: | :: | :: | :: | :: | :: | :: | :: | :: | :: | :: | :: | :: | :: | :: | :: | :: | :: | :: | :: | :: | :: | :: | :: | :: | :: | :: | :: | :: | :: | :: | :: | :: | :: | :: | :: | :: | :: | :: | :: | :: | :: | :: | :: | :: | :: | :: | :: | :: | :: | :: | :: | :: | :: | :: | :: | :: | :: | :: | :: | :: | :: | :: | :: | :: | :: | :: | :: | :: | :: | :: | :: | :: | :: | :: | :: | :: | :: | :: | :: | :: | :: | :: | :: | :: | :: | :: | :: | :: | :: | :: | :: | :: | :: | :: | :: | :: | :: | :: | :: | :: | :: | :: | :: | :: | :: | :: | :: | :: | :: | :: | :: | :: | :: | :: | :: | :: | :: | :: | :: | :: | :: | :: | :: | :: | :: | :: | :: | :: | :: | :: | :: | :: | :: | :: | :: | :: | :: | :: | :: | :: | :: | :: | :: | :: | :: | :: | :: | :: | :: | :: | :: | :: | :: | :: | :: 
564 SHLEEVGKLANKYNHRLTMHPGQYTQIASPREVVVDSAIRDL-----AYHDEILSRMK 616
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            Truncated fusion proteins and related polynucleotides useful for screening for compounds which modulate Weel function
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             GST:Weel(11-72); glutathione-S-transferase; GST; WEEl; human; mitosis; cell cycle checkpoint; cyclin-dependent kinase; Cdsl.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         748 RVFDFPPCDPTMDLMIEAKEKEQAVFELCRRYELQNPPCPLEIMGPEYD 796
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            : | || |: :||: : | || 827 ligwvphcdtlhalirdyrekkillniehrimlr------mapdyd 867
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  Glutathione-S-transferase:Weel(11-72) fusion proein.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      Claim 5; Page 37-38; 46pp; English.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           AAY31953 standard; Protein; 310 AA.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       Russell P, Boddy MN, Funari B;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   Chimeric - Schistosoma mansoni.
Chimeric - Homo sapiens.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                99WO-US06704.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            98US-0079752.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  21-DEC-1999 (first entry)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            (SCRI ) SCRIPPS RES INST.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               WPI; 1999-591103/50.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         N-PSDB; AAZ20066.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     WO9949061-A1.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      731 -----
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            27-MAR-1998;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                26-MAR-1999;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                30-SEP-1999
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               AAY31953;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         AAY31953
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          g
                                                                                                                                                                   ò
                                                                                                                                                                                                                                                                                                                           ö
                                                                                                                                                                                                                                                                                                                                                                                                  g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      οy
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              QΥ
```

```
242 glrrsqrsmnlnra---tll-------apptpsslydannst 273
                                                                                                                                                                                                                       125 LSKLPEMLKMFEDRLCHKTYLNGDHVTHPDFMLYDALDVVLYMDPMCLDAFPKLVCFKKR 184
                                                                                                                                                                                                                                                                                     231 -LRLLKRNIQISKRIVFTILKQKAFKGNHPCVPSVCTITYSRFHCLPDTLKSLLPMSSKT 289
                                                                                                                                                                                                         65 VKLTQSMAIIRYIADKHNMLGGCPKERAEISMLEGAVLDIRYGVSRIAYSKDFETLKVDF 124
                                                                                                         63; Gaps
                                                                                                                                                        5 PILGYWKIKGLVQPTRLLLEYLEEKYEEHLYERDEGDKWRNKKFELGLEFPNLPYYIDGD 64
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        AAR84565-R84569 are polypeptides of the TCR27 protein of T.cruzi
The proteins are all fusion products with glutathione-S-transferase
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           New polypeptide(s) from the TCR27 protein of Trypanosoma cruzi - a immunoassay reagent for specific diagnosis of Chagas disease, also related nucleic acid and transformed cells
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     TCR27; Chagas disease; repeat unit; diagnosis; blood screening; recombinant; fusion protein; glutathione-5-transferase.
                                                                          DB 20; Length 310;
                                                                                                                                                                                                                                                                                                                                     185 IEAIPQIDKYLKSSKYIAWPLQGWQATFGGGDHPPKSDHLVPRGSM-----
                                                                                                                                                                                                                                                                                                                                                                                                                                                               290 TLSMLPQVNIGANSFSAETPVDLKKENETELANISGPHKKSTSTSTRKRA 339
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             274 -----stssgkp----ntsftsligprkqttsspsfsha 304
                                                                          28.0%; Score 1216.5; DB 20; Length
69.7%; Pred. No. 8.6e-84;
tive 18; Mismatches 25; Indels
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        Trypanosoma cruzi TCR27 N-terminal polypeptide, Ag2-2.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           Disclosure; Page 33-34; 68pp; English.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             AAR84566 standard; Protein; 354 AA.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       95WO-US03191.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            09-MAY-1996 (first entry)
                                                                                                            Matches 244; Conservative
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  (KIRC/) KIRCHHOFF L V.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  WPI; 1995-344618/44.
                                                                             Query Match
Best Local Similarity
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            Trypanosoma cruzi.
                                310 AA;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 N-PSDB; AAT05330.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     (OTSU/) OTSU K.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  Kirchhoff LV,
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     20-MAR-1995;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    24-MAR-1994;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       28-SEP-1995.
                                Sequence
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             AAR84566;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   AAR84566
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              qq
S X C
                                                                                                                                                                                                                                                                                                      qq
```

```
7;
(GST) and some contain a linker sequence. The TCR27 protein comprises a 95 amino acid (aa) N terminal region; 69 repeats of a highly conserved 14 as a sequence and a 68 ara C-terminal region. This sequence encodes the GST sequence and the TCR27 N-terminal polypeptide Ag2-2. The TCR27 bolypeptides of the invention are useful for the diagnosis of Chagas disease (American Trypanosomiasis), they are capable of detecting anti-T.cruzl antibodies, or for blood screening. The TCR27 protein has epitopes to which most T.cruzl infected individuals have antibodies. The TCR27 polypeptides will not react with serum from patients with leishmanlasis, schistosomiasis, or autoimmune disease and are hence less likely to cause false positives in diagnosis.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              185 IBAIPQIDKYLKSSKYIAWPLQGWQATFGGGDHPPKSDHLVPRG--SMLRLLKRNIQISK 242
                                                                                                                                                                                                                                                                                                                                                                                                                     125 LSKLPEMLKMFEDRLCHKTYLNGDHVTHPDFMLYDALDVVLYMDPMCLDAFPKLVCFKKR 184
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           243 RIVFTILKOKAFKGNHPCVPSVCTIT-----YSRFHCLPD----TLKSLLPMSSKT 289
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               242 k-emtkirekvkkaekekidainratkleeernqaykaahkaeeekaktfqrlitfese- 299
                                                                                                                                                                                                                                                                                                                                                                                                       65 VKLTQSMAIIRYIADKHNMLGGCPKERAEISMLEGAVLDIRYGVSRIAYSKDFETLKVDF 124
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    290 TLSMLPQVNIGANSFSAETPVDLKKENETELANISGPHKKSTSTSTRKRARSSKKKATDS 349
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          27; Gaps
                                                                                                                                                                                                                                                                                                                                   5 PILGYWKIKGLVQPTRLLLEYLEEKYEEHLYERDEGDKWRNKKFELGLEFPNLPYYIDGD 64
                                                                                                                                                                                                                                                                                                                                                     3 pilyywkikglvqptrllleyleekyeehlyerdegdkwrnkkfelglefpnlpyyidgd 62
                                                                                                                                                                                                                                                               epithelial cell; growth factor receptor; interleukin-3; IL-3; erythropoletin: EPO; negative regulator; signal transduction; modulation; activity; inhibitor; stimulation; cytokine therapy; antisense therapy; gene therapy; treatment; diagnosis;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       Haematopoietic cell phosphatase; HCP; cloning; expression;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              AAY52288 standard; Protein; 824 AA.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           glutathione-S-transferase; fusion.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    Human HCP/GST fusion protein.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         99WO-US08228
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    09-FEB-2000 (first entry)
                                                                                                                                                                                                                                                                                                Matches 248; Conservative
                                                                                                                                                                                                                                                                                Similarity
                                                                                                                                                                                                              354 AA;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  Homo sapiens.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         15-APR-1999;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   W09954450-A1
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      28-OCT-1999
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         350 V 350
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          351 i 351
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  Synthetic.
                                                                                                                                                                                                                Sequence
                                                                                                                                                                                                                                                                 Query Match
Best Local 9
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   AAY 52288;
     8
                                                                                                                                                                                                                                                                                                                                                                    a
                                                                                                                                                                                                                                                                                                                                                                                                         ò
                                                                                                                                                                                                                                                                                                                                                                                                                                        g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             a
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       ογ
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  ò
```

```
CC has been shown to accluded protein found primarily in hasmatopoietic cells, and to some extent in epithelial cells. It has been shown to associated protein found primarily in hasmatopoietic cells, and to some extent in epithelial cells. It following ligand stimulation and acts as a negative regulator collowing ligand stimulation and acts as a negative regulators. The invention relates to a novel contropoietin (EPO) receptors. The invention relates to a novel method of identifying compounds that modulate HCP activity. This comprises combining a modulator of HCP protein activity with HCP protein and HCP protein. Inhibitors of the modulator on the activity of the HCP protein. Inhibitors of HCP can be used to be a negative regulator of one or more signal transduction pathways in hasmatopoietic cells, and the identification of a HCP inhibitor would be covide a synthetic stimulator to the hasmatopoietic system could be provide a synthetic stimulator to the hasmatopoietic system could be consent on the identification of a HCP inhibitor may also therefore a sheartopoietic cells and it is possible that a HCP inhibitor may also function as a hasmatopoietic cell potentiator, enhancing the effect of growth factors, decreasing the necessary dose of the factor. The HCP polynucleotides are sources of probes and primers, and can be used to design antisense sequences, and in gene therapy. The protein, or active fragments thereof, may be used to treat or diagnose disorders which require the modulation of HCP consider the effect of growth require the modulation of HCP considered to the effect of a child the action of HCP considered to the effect of polynucleotides are sources, and in gene therapy. The protein, or active fragments thereof, may be used to reactive fragments thereof, may be used to receive the modulation of HCP considered to the effect of polynucleotides are sources, and in gene therapy. The protein of the considered inhibit the action of HCP considered to the HCP activity. Small modes the considered in the HCP c
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          65 VKLTQSMAIIRYIADKHNMLGGCPKERAEISMLEGAVLDIRYGVSRIAYSKDFETLKVDF 124
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          125 LSKLPEMLKMFEDRLCHKTYLNGDHVTHPDFMLYDALDVVLYMDPMCLDAFPKLVCFKKR 184
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       185 IEAIPQIDKYLKSSKYIAWPLQGWQATFGGGDHPPKSDHLVPRGS----MLRLLKRNIQIS 241
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  Gaps
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       5 PILGYWKIKGLVQPTRLLLEYLEEKYEEHLYERDEGDKWRNKKFELGLEFPNLPYYIDGD 64
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     inhibit the action of HCP can also augment the effect of EPO, which meaning that a lower therapeutic dose of EPO may be utilised.
                                                                                                                                                                                                                                             Assay methods for modulators of hematopoietic cell phosphatase
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  9
                                                                                                                                                                                                                                                                                                                                                  This sequence represents a full-length human haematopoietic cell phosphatase (HCP)/glutathione-5-transferase (GST) fusion protein. This was expressed in E. coli and subsequently purified for use in a variety of activity assays.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       27.9%; Score 1212; DB 21; Length 824; 89.3%; Pred. No. 7.7e-83;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               Indels
                                                                                                                             Xu D;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  Pred. No. 7.7e
6; Mismatches
                                                                                                                     Jolliffe L, Barbone F, Shanahan M,
                                                                                                                                                                                                                                                                                               Claim 3; Fig 4B; 57pp; English.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          242 KRIVFTILKQKAFKGNHPCVPS 263
                                                                     (ORTH ) ORTHO-MCNEIL PHARM INC.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       240 gldaetlikgrgvhgsflarps 261
                     98US-0082072
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     Matches 234; Conservative
                                                                                                                                                                          WPI; 2000-013244/01
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               Best Local Similarity
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               824 AA;
                                                                                                                                                                                                     N-PSDB; AAZ32916
                       17-APR-1998;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               Sequence
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          Query Match
òλ
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             qq
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               Óγ
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             Q
Q
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  ŏ
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     Q
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     ŏ
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    q
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          õ
```

```
546 iaapavffkseaivhpdqvvqyiggdecqigynplqmallwntlatrevnllhqaltyrh 605
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      |:| : | : | : | 654 ---fdgsfargvpfgynpstgdcrvsgtaaalvglagddp--havdrikllysi---- 702
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  688 DWHHHNIVPGTLREGSLDLMPL---IPTIRE-----TWTRKGITQKQHYSES 731
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 703 -----alstgglpliylgdevgtlndddwsqdsnksddsrwahrprynealyagr 752
63 vkltgsmaiiryiadkhumlggcpkeraeismlegavldirygvsriayskdfetlkvdf 122
                                                         125 LSKLPEMLKMFEDRLCHKTYLNGDHVTHPDFMLYDALDVVLYMDPMCLDAFPKLVCFKKR 184
                                                                                     185 IEAIPQIDKYLKSSKYIAWPLOGWQATFGGGDHPPKSDHLVPRGS----MLRLLKRNI-- 238
                                                                                                                                                                                                                                 183 ieaipqidkylksskyiawplqgwqatfgggdhppksd-lvprgspefmilqylktrild 241
                                                                                                                                                                                                                                                                                        239 ------QISKRIVFTILKQKAFKGNHPCVPSVCTITYSRFHCLPDTLKS 281
                                                                                                                                                                                                                                                                                                                                                                                                         282 LLPMSSKTTLSMLPQVNIGANSFSAETPVDLKKENETE--LAN--ISGPHKKSTSTSTRK 337
                                                                                                                                                                                                                                                                                                                                                                                                                                                      292 llaqawqs----ysqrn-----sslkdidiarennpdwilsnkqvggvcyvdlfagdlk 341
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   342 ------glkdkipyfgelgltylhlmplfkcpegksdggyavssyrdvnpalgti 390
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       387 -----SEEQISKATKKVKQKEEEEYVEEVDEKSLKNESSSDEFEPVVPEQLETPI 436
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           391 gdlreviaalheagisavvdfifnhtsneh--ewaqrcaagdplfdnfyyifpdr-rmpd 447
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               448 qydrt.lreifpdghpggfsgled-------grwvwttfnsfqwdlnysnp 490
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        487 -ILRSMKERVFCSRTCRITTIQRDGLESV-KQLGT--QNVLDLIKLVE-WNHNFGIHFMR 541
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   542 VSSDLFPFASHA-------KYGY----TLEFAQSHLEEVGKL--ANKYNH 578
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    579 RLTMHPGQYTQIASPREVVVDSAIRDLAY-----HDEILSRMKLNEQLNKDAVLIIHL 631
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        606 nlpehtawvnyvrshddigwtfadedaaylgisgydhrqflnrffvnr------ 653
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 632 GGTFEGKKETLDRFRKNYQ----RLSDSVKARLVLENDDVSWSVQDLLPLCQELNIPLVL 687
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 Syphilis; G15 antigen; membrane antigen; glutathione-S-transferase; GST; fusion protein; membrane antigen; diagnosis.
                                                                                                                                                                                                                                                                                                                            242 iytpeqragieksedwrqfsrrm-----dthfpklmneldsvygnneallpmlem
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         338 RARSSKKKATDSVSDKI---DESVASYDSSTHLRRSSRSKKPVNYNSSSESE---
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           437 SKRRRSRSSAKNLEKESTMNLJDHAPREMFDCLDKPIPWRGRLGYACLNT---
                                                                                                                                                                                                         Location/Qualifiers
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         AAR79948 standard; Protein; 362 AA.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  Chimeric unidentified bacterium.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          Chimeric Treponema pallidum;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             (first entry)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      732 ADPTAISG 739
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                753 ndpstaag 760
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          26-FEB-1996
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   AAR79948;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             Protein
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              AAR79948
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         δλ
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                XX
FH
FT
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 qq
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              qq
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  qq
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               qq
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            δ
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                qq
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             óγ
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           δŏ
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        ŏλ
                                                                                                                                                                                                                                       QQ
                                                                                                                                                                                                                                                                                                                                                     qq
                                                                                                                                                                                                                                                                                                                                                                                                                  δy
                                                                                                                                                                                                                                                                                                                                                                                                                                                                     q
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               ŏ
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      qq
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             δy
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 qq
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              ŏ
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           ó
                                                               δy
                                                                                                                                                                             δ
                                                                                                                                                                                                                                                                                                 ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          This invention describes a novel amylosucrase (AS), immobilizable on a solid phase, which comprises a fusion protein (FP) of functional units of AS, an anchoring sequence, and optionally additional auxiliary sequences. The invention also describes (1) nucleic acid (1) encoding FP; (2) expression vector containing (1) and able to express FP in a host cell; Escherichia coli containing (1) and able to express FP in a nucleotides that encodes an epitope, a high-affinity binding partner or SST (glutathione-S-transferase); (4) solid phase for immobilizing AS comprising FP immobilized on Sepharose for production of poly(1,4-alpha-glucan) (II); and (6) biocatalytic production of (II).

SS is used in production of poly(1,4-alpha-glucans) (II), useful for producing films, as food additives, as starting materials for producing for
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          29;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      Cyclodextrins and as auxiliaries in pharmaceutical formulations. Immobilized As makes possible efficient, inexpensive and continuous production of poly(1,4-alpha-glucans) (I), and it can be used repeatedly. Compared with known methods, specificity is improved (increased yield of (II) and reduced formation of palatinose) and reaction is complete within 24 hours, compared to 48-72 hours for batch methods. This sequence represents the expression vector construct pGEX-4T-1-AmsUs which contains the amylosucrase (EC 2.4.1.4) described in
                                                                                                                                                                                                                                                                                                 Amylosucrase; EC 2.4.1.4; fusion protein; GST; glutathione-S-transferase; poly(1,4-alpha-glucan); film production; food additive; cyclodextrin.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              Amylosucrase\ immobilized\ as\ fusion\ protein\ with\ anchoring\ sequence,\ useful\ in\ continuous\ preparation\ of\ poly(1,4-alpha-glucan)\ from\ sucrose
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  65 VKLTQSMAIIRYIADKHNMLGGCPKERAEISMLEGAVLDIRYGVSRIAYSKDFETLKVDF 124
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          239; Indels 203; Gaps
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   3 pilgywkikglvqptr111leyleekyeehlyerdegdkwrnkkfelglefpnlpyyidgd 62
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 5 PILGYWKIKGLVQPTRLLLEYLEEKYEEHLYERDEGDKWRNKKFELGLEFPNLPYYIDGD 64
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       DB 22; Length 856;
                                                                                                                                                                                                                                             Expression vector pGEX-4T-1-AmSu5 containing amylosucrase.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    27.8%; Score 1208.5; DB 39.0%; Pred. No. 1.5e-82;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        Polakowski T, Held A, Gallert K;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             75; Mismatches
                                                                   AAB70883 standard; Protein; 856 AA.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               Claim 2; Page 34-38; 38pp; German.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 04-OCT-2000; 2000WO-EP09695.
                                                                                                                                                                                      12-JUL-2001 (first entry)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          Conservative
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               WPI; 2001-328330/34.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       Query Match
Best Local Similarity
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               (AXIV-) AXIVA GMBH.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 856 AA;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           N-PSDB; AAF61712
                                                                                                                                                                                                                                                                                                                                                                                                                                                WO200125449-A2.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              invention
                                                                                                                                                                                                                                                                                                                                                                                            Unidentified.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        07-OCT-1999;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          12-APR-2001.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          Matches 331;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       Sequence
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        Bengs H,
                                                                                                                            AAB70883;
                                    AAB70883
            RESULT
```

δ g ŏ 13-JUN-1996 (first entry)

```
123 lsklpemlkmfedrlchktylngdhvthpdfmlydaldvvlymdpmcldafpklvcfkkr 182
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       185 IEAIPQIDKYLKSSKYIAWPLQGWQATFGGGDHPPKSDHLVPRGSML----RLLKRNIQI 240
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                298 NIGANSFSAETPVD-----LKKENETELANIS-----GPHKKSTSTSTRKRARSSKKKA 346
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  325
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          LSKLPEMLKMFEDRLCHKTYLNGDHVTHPDFMLYDALDVVLYMDPMCLDAFPKLVCFKKR 184
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   241 SKRIVFTILKQKAFKGNHP---CVPSVCTITYSRFHCLPDTLKSLLPMSSKTTLSMLPQV 297
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               3 pilgywkikglvqptrllleyleekyeehlyerdegdkwrnkkfelglefpnlpyyidgd 62
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            65 VKLTQSMAIIRYIADKHNMLGGCPKERAEISMLEGAVLDIRYGVSRIAYSKDFETLKVDF 124
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     64
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    69; Gaps
                                                                                                                                                                                                                                                                                                                                                                        G17 antigen (AAR79948) is a fusion protein composed of glutathione-S-transferase and the 17 kDa membrane antigen of Treponema pallidum. It is obtd. by insertion of a DNA fragment coding for the 17 kDa antigen into vector pWG6A and expression in E. coli. G17 and similarly produced G15 (AAR79947) are used for diagnosis of syphilis.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      5 PILGYWKIKGLVQPTRLLLEYLEEKYEEHLYERDEGDKWRNKKFELGLEFPNLPYYIDGD
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  27.8%; Score 1207.5; DB 16; Length 362; 61.6%; Pred. No. 5.1e-83; Live 24; Mismatches 65; Indels 69;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               283 elalekksapspltyrgtwmvredgivelslyssegskaphek------
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         326 --elyelidsnsvrymgapgagkpskemapf------yvlkktkk 362
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            347 TDSVSDKIDESVASYDSSTHLRRSSRSKKPVNYNSSSESESEEQISKATKK 397
                                                                                                                                                                                                                                                                                               New assay for anti-Treponema pallidum antibody - useful for diagnosis of syphilis.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  242 kaekvecalkggifrgtlpaadcpgidttvtfn-------
                                                                                                                                                                                                                                      <u>с</u>і
                                                                                                                                                                                                                                      Ueno
                                                                                                                                                                                                                                                                                                                                           Disclosure; Page 10-11; 18pp; English.
              229..362
/label= 17k_antigen
                                                                                                                                                                                                                                      Okada M,
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        AAR87091 standard; Protein; 426
                                                                                                                          95EP-0400423
                                                                                                                                                      94JP-0054672
 /label= GST
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 Matches 253; Conservative
                                                                                                                                                                                     (FJRE ) FUJI REBIO INC (FJRE ) FUJI REBIO KK.
                                                                                                                                                                                                                                  Fujii N, Fujimura K,
                                                                                                                                                                                                                                                                WPI; 1995-303656/40.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    Similarity
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      362 AA;
                                                                                                                        28-FEB-1995;
                                                                                                                                                      28-FEB-1994;
                                                            EP670494-A2
                                                                                           06-SEP-1995
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       Sequence
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     Query Match
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  Best Local
                Protein
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         125
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         AAR87091
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         RESULT
ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        q
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            셤
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              ò
```

```
Sequence
                                   Query Match
δ
                                        В
                                          δ
                                            δy
                                              qq
                                                δλ
                           .;
8
```

```
This is the sequence of a glutathione-S-transferase/turkey prolactin fusion protein. The fusion protein can be used to actively immunise birds, esp. turkeys, pref. during the rearing stage of the animal's life, prior to sexual maturity, norder to prevent or treat broodiness in the birds. Alternatively the prolactin protein (R879090) can be used to raise antibodies for passive immunisation, for preventing or
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   65 VKLTQSMAIIRYIADKHNMLGGCPKERAEISMLEGAVLDIRYGVSRIAYSKDFETLKVDF 124
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        125 LSKLPEMLKMFEDRLCHKTYLNGDHVTHPDFMLYDALDVVLYMDPMCLDAFPKLVCFKKR 184
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   62
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     Gaps
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   64
                                                                                                                                                                                                                                                                                     /note= "turkey prolactin portion of fusion protein"
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                Controlling broodiness in birds, esp. turkeys - by active immunisation with prolactin fusion protein or passive immunisation
                                                Turkey; prolactin; glutathione-S-transferase; fusion protein; egg; immunisation; bird; turkey; broodiness; antibody.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         3 pilgywkikglvqptrllleyleekyeehlyerdegdkwrnkkfelglefpnlpyyidgd
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  5 PILGYWKIKGLVQPTRLLLEYLEEKYEEHLYERDEGDKWRNKKFELGLEFPNLPYYIDGD
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                4 ;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          Length 426;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                Indels
                                                                                                                                                                                                                             /note= "GST portion of fusion protein" 228..426
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              treating broodiness during the egg laying stages.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       27.7%; Score 1206; DB 16;
95.0%; Pred. No. 8.4e-83;
tive 2; Mismatches 6;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                (INRG ) INRA INST NAT RECH AGRONOMIQUE
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   Zadworny D;
Turkey prolactin/GST fusion protein.
                                                                                                                                                                                  Location/Qualifiers
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 Claim 3; Page 29-31; 41pp; French.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   with anti-prolactin antibody
                                                                                                                                                                                                                                                                                                                                                                                                                                          94FR-0005550
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             94FR-0005550
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          Matches 228; Conservative
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      Karatzas C,
                                                                                                                                                                                                             1..227
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    WPI; 1995-384532/50.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     Local Similarity
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 426 AA;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            N-PSDB; AAT03457
                                                                                                                                                                                                          Misc_feature
                                                                                                                                                                                                                                                          Misc-Feature
                                                                                                                                                                                                                                                                                                                                                                                                                                          05-MAY-1994;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             05-MAY-1994;
                                                                                                                                                                                                                                                                                                                                   FR2719480-A1
                                                                                                                                                                                                                                                                                                                                                                                      10-NOV-1995
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      Guemene D,
                                                                                                                              Synthetic
```

5;

Search completed: January 15, 2002, 13:49:55 Job time: 284 sec

AAR87091;

?

```
January 15, 2002, 14:06:09; Search time 134.53 Seconds (without alignments) 651.283 Million cell updates/sec
                                                                                                                                                                                                                                                                                                 1 MLRLLKRNIQISKRIVFTIL.......KRLTARKRRSRKEEVEEDEK 599
                            Compugen Ltd.
GenCore version 4.5
Copyright (c) 1993 - 2000 Comp
                                                                                                                                                                                                                                                    US-09-724-296-2_COPY_230_828
                                                                                                 OM protein - protein search, using sw model
                                                                                                                                                                                                                                                                            Perfect score:
                                                                                                                                                                                                                                                                                                      Sednence:
                                                                                                                                                Run on:
```

473505 Total number of hits satisfying chosen parameters: 473505 seqs, 146272329 residues BLOSUM62 Gapop 10.0 , Gapext 0.5 Searched:

Scoring table:

Minimum DB seq length: 0 Maximum DB seq length: 200000000

Listing first 45 summaries Maximum Match 100% Post-processing: Minimum Match 0%

sp_unclassified:* sp_human:*
sp_invertebrate:*
sp_mammal:* sp_virus:* sp_vertebrate:* sp_organelle:* sp_archea:* sp_bacteria:* sp_rodent:* sp_plant:* SPTREMBL 17:* sp_phage:* sp_fungi:* sp_mhc:* 10: Database :

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

	1 1 4 4 4 4 4 4 4 4 4 4 4 4 4 4 4 4 4 4	Description	O10988 schizosacch							O9esk9 musculu	O45198 caenorhabdi	Ogarz3 caenorhabdi	O26104 plasmodium	Q9u414 plasmodium	006704 saccharomyc	002061 caenorhabdi	09u7e0 caenorhabdi			Q9u430 plasmodium	004010 onchocerca
	9	TD	Q10988	001408	O9L4F0	09L4E9	Q9K9P8	6M0S60	O9RTE6	09ESK9	045198	Q9GRZ3	026104	Q9U414	006704	002061	Q9U7E0	026007	013885	Q9U430	004010
	9	9 :	m	e	2	7	7	7	7	11	S	Ŋ	Ŋ	Ŋ	٣	2	S	5	m	S	Ŋ
	Query		599	929	317	325	322	305	326	1588	368	1641	782	782	911	1359	1359	782	401	782	1957
æ	Query	ייייייייייייייייייייייייייייייייייייייי	100.0	32.5	11.5	11.5	8.6	7.4	7.4	5.3	5.0	5.0	2.0	4.8	4.8	4.8	4.8	4.7	4.7	4.7	4.7
	Score	2000	3112	1011.5	359	359	306.5	229.5	229.5	164	156.5	155	154.5	149.5	148.5	148.5	148	147.5	145.5	145.5	145
	Result		1	7	m	4	5	9	7	œ	6	10	11	12	13	14	15	16	17	18	19

025730 plasmodium				Q9yhd5 rana catesb	Q22342 caenorhabdi	034097 streptococc	Q9c3y7 candida alb	Q69513 human herpe	O54875 rattus norv	Q9qyx6 mus musculu	O9qyx7 mus musculu	Q9qyz0 strongyloce	Q9u8g1 plasmodium		Q9vqn4 drosophila	Q40363 medicago sa	Q9vcp0 drosophila	001900 caenorhabdi	Q9u429 plasmodium	O56300 human herpe	Q91749 xenopus lae	Q911h8 arabidopsis	Q9u431 plasmodium		Q9bzf9 homo sapien
025730	025875	026026	075130	3 Q9YHD5	022342	034097	Q9C3Y7		1 054875		1 Q9QYX7	09GYZ0	Q9U8G1	096124	Q9VGN4	040363	Q9VCP0	001900	090429		3 091749	0 09гінв	090431	Q9HCL1	Q9BZF9
2	2	5 5	5	5 13	0 5	3	7 3	П	2 11	П	П	3	. 5	. 5	9	5 10	3	. 5	2	5 12			2	1 4	5
782	78.	816	84	826	133(65.	1197	120	1702	4833	5038	1463	1.680		399					1205	480	61(782	1151	1416
4.6	4.6	4.6	4.6	4.5	4.5		4.5	4.5	4.5	4.5	4.5	4.5	4.5	4.5	4.5	4.5	4.5	4.5	4.5	4 . 4	4.4	4.4	4.4	4 . 4	4.4
144.5	144.5	144.5	144	141.5	141.5	141	141	14.1	141	141	141	140.5	140.5	140.5	140	139.5	139.5	13	138.5	m	137.5	137.5	137.5	137.5	137.5
20	21	22	23		25	56	27	28	29	30	31	32	33	34	35	36	37	38	39	40	41	42	43	44	45

ALIGNMENTS

```
Takao M., Yonemasu R., Yamamoto K., Yasui A.;
Characterization of a UV endonuclaase gene from the fission yeast
Schizosaccharomyces pombe and its bacterial homolog.";
Nucleic Acids Res. 24:1267-1271(1996).
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           Lyne M., Rajandream M.A., Barrell B.G., Lucas M., Gaillardin C.; Submitted (JUN-1998) to the EMBL/GenBank/DDBJ databases. EMBL: 078571; BAA11415-1; -. EMBL: U78487; AAC49664-1; -. EMBL; AL023859; CAA19577.1; -.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           (2)
SEQUENCE FROM N.A.
MEDLINE=97175806; PubMed=9023111;
Davey S., Nass M.L., Ferrer J.V., Sidik K., Eisenberger A., Mitchell D.L., Freyer G.A.:
"The fission yeast UVDR DNA repair pathway is inducible.";
Nucleic Acids Res. 25:1002-1008(1997).
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             Endonuclease.
SEOUENCE 599 AA; 68816 MW; C09ED7219D5895BE CRC64;
                                                                                                                                                                                                                                                                                                                                                  Schizosaccharomyces pombe (Fission yeast).
Waryota; Fungi, Ascomycota; Schizosaccharomycetes; Schizosaccharomycetales; Schizosaccharomyces.
                                                                                                                                                                                                  Last sequence update)
Last annotation update)
                                                                              599 AA
                                                                                                                                                             Created)
                                                                              PRT;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        MEDLINE=96188860; PubMed=8614629;
                                                                                                                                                    01-NOV-1996 (TrEMBLrel. 01, 01-NOV-1996 (TrEMBLrel. 01, 01-NOV-1998 (TrEMBLrel. 08,
                                                                              PRELIMINARY;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                [3]
SEQUENCE FROM N.A.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  SEQUENCE FROM N.A.
                                                                                                                                                                                                                                                                            UV-ENDONUCLEASE.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    NCBI_TaxID=4896;
                                                                                                               Q10988; P87339;
                                                                                                                                                                                                                                                                                                                 UVDE OR UVE1+.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              STRAIN=972H-;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   STRAIN=SP972
                                                                          010988
RESULT
                                 010988
                                                                                  HDDDALAND REPRETED DO THE STREET OF THE STRE
```

Query Match

DB 3; Length 599; 100.0%; Score 3112;

```
0;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                "A eukaryotic gene encoding an endonuclease that specifically repairs DNA damaged by ultraviolet light."; EMBO J. 14:2393-2399(1995).
                                                                                                                                                                                                                                                                                                                                                                            EWNHNFGIHFMRVSSDLFPFASHAKYGYTLEFAQSHLEEVGKLANKYNHRLTMHPGQYTQ 360
                                                                                                                                                                                                                                                                                                                                                                                                                       361 IASPREVVVDSAIRDLAYHDEILSRMKLNEQLNKDAVLIIHLGGTFEGKKETLDRFRKNY 420
                                                                                                                                                                                                                                                                                                                                                                                                                                       QRLSDSVKARLVLENDDVSWSVQDLLPLCQELNIPLVLDWHHHNIVPGTLREGSLDLMPL 480
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               481 IPTIRETWTRKGITQKQHYSESADPTAISGMKRRAHSDRVFDFPPCDPTMDLMIEAKEKE 540
                                                                                                                           TLSMLPQVNIGANSFSAETPVDLKKENETELANISGPHKKSTSTSTRRRARSSKKKATDS 120
                                                                                                                                                                         121 VSDKIDESVASYDSSTHLRRSSRSKKPVNYNSSSESESEGISKATKKVKQKEEEEYVEE 180
                                                                                                                                                                                                     121 VSDKIDESVASYDSSTHLRRSSRSKKPVNYNSSSESEBERGISKATKKVKQKEBEBEYVEE 180
                                                                                                                                                                                                                                    181 VDEKSLKNESSSDEFEPVVPEQLETPISKRRRSRSSAKNLEKESTMNLDDHAPREMFDCL 240
                                                                                                                                                                                                                                                                                               DKPIPWRGRLGYACLNTILRSMKERVFCSRTCRITTIQRDGLESVKQLGTQNVLDLIKLV 300
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                421 QRLSDSVKARLVLENDDVSWSVQDLLPLCQELNIPLVLDWHHHNIVPGTLREGSLDLMPL 480
                                                                                                          61 TLSMLPQVNIGANSFSAETPVDLKKENETELANISGPHKKSTSTSTRKRARSSKKKATDS 120
                                                            1 MLRLLKRNIQISKRIVFTILKQKAFKGNHPCVPSVCTITYSRFHCLPDTLKSLLPMSSKT 60
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        541 QAVFELCRRYELQNPPCPLEIMGPEYDQTRDGYYPPGAEKRLTARKRRSRKEEVEEDEK 599
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         Yajima H., Takao M., Yasuhira S., Zhao J.H., Ishii C., Inoue H.,
                0;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           Eukaryota; Fungi; Ascomycota; Pezizomycotina; Sordariomycetes;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        Score 1011.5; DB 3; Length 656; Pred. No. 2.2e-56;
                  Indels
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             656 AA; 74413 MW; 4B231E7A8735C86F CRC64;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             01-NOV-1996 (TrEMBLrel. 01, Created)
01-MAY-1999 (TrEMBLrel. 10, Last sequence update)
01-MAY-1999 (TrEMBLrel. 10, Last annotation update)
Pred. No. 7.9e-190;
                  Mismatches
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        Sordariaceae; Neurospora
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    MEDLINE=95292980; PubMed=7774597;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   PRT;
                0;
   100.08;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          32.5%;
36.6%;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            EMBL; D11392; CAB20113.1;
                  Conservative
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 PRELIMINARY;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          Query Match
Best Local Similarity
   Similarity
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            Neurospora crassa.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    SEQUENCE FROM N.A.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       STRAIN=74-OR23-1A;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         NCBI_TaxID=5141;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             UV - ENDONUCLEASE.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           Sordariales;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              Endonuclease
                  599;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     Yasui A.;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             SEQUENCE
 Best Local
Matches 59
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             001408;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 001408
                                                                                                                                                                                                                                                                                                                                                                                                                                                      361
                                                                                                                                                                                                                                                                                                                                                                                         301
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    421
                                                                                                                                                                                                                                                                                                                                                          301
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  RESULT
                                                                                                              οy
                                                                                                                                         qq
                                                                                                                                                                                                     엄
                                                                                                                                                                                                                                    δy
                                                                                                                                                                                                                                                                QQ
                                                                                                                                                                                                                                                                                               Qγ
                                                                                                                                                                                                                                                                                                                           qq
                                                                                                                                                                                                                                                                                                                                                          δý
                                                                                                                                                                                                                                                                                                                                                                                      qq
                                                                                                                                                                                                                                                                                                                                                                                                                         δ
                                                                                                                                                                                                                                                                                                                                                                                                                                                   g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      qq
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 ö
```

```
18;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             115 GIPVPSVVSRF-PTAPYHHKSTNAEEREAKEPVLKTHSKDVBREAEIGVDDVVKMEPAAT 173
                                                                                                                                                                                                                                 427 VKARLVLENDDVSWSVQDLLPLCQELNIPLVLDWHHHNIV--PGTLREGSLDLM--PLIP 482
                                                                                                                                                                                                                                                                                                                                                                                                                                                                           483 TIRETWIRKGITOKOHYSESADPTAISGMKRRAHSDRVFDFPPCDPTMDLMIEAKEKEQA 542
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        ---LEIMGPE--- 565
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            533 VFELMRTFKLPGFEKINDMVPYDRDDENRPAPPVKAPKKKGGKRKRTTDEEAAEPEEVD 592
                                                                                                                     63 AGRMMGKDANGHCLREGKEQEEGVKMAIEGLARMERRLQRATKRQKKQLEED----- 114
                                                                                                                                                     185 SLKNESSSDEFEPVVPEQLETPISKRRRS----RSSAKNLEKESTMNLDDHA----- 232
                                                                                                                                                                                                               233 ----PREMFDCLDK------PIPWRGRLGYACLNTILRSMKERVFCSRTCRIT 275
                                                                                                                                                                                                                                                                         ---ORDGLESVKQLGTQNVLDLIKLVEWNHNFG 307
                                                                                                                                                                                                                                                                                                        234 SIVDHRHPLQFEDEPEHHLKNKPDKSKEPQDELGHKFVQELGLANARDIVKMLCWNEKYG 293
                                                                                                                                                                                                                                                                                                                                     308 IHFMRVSSDLFPFASHAKYGYTL-EFAQSHLEEVGKLANKYNHRLTMHPGQYTQIASPRE 366
                                                                                                                                                                                                                                                                                                                                                                                               367 VVVDSAIRDLAYHDEILSRMKLNEQLNKDAVLIIHLGGTFEGKKETLDRFRKNYQRLSDS 426
                              110 ARSSKKKATDSVSDKIDESVASYDSST----HLRRSSR-----SKKPVNYNSSSESE- 157
                                                                                          --SEEQISKATKKVKQKEEEEYVEEVDEK 184
88; Mismatches 149; Indels 181; Gaps
                                                   -----YYPPGAEKRLTARKRRSRK----EEVEED 597
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       593 TAADDVKDAPEGPKEVPEEERAMGGPYNRVYWPLGCEEWLKPKKREVKKGKVPEEVEDE 651
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     317 AA; 36990 MW; A6AC35F5800A22C3 CRC64;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                01-OCT-2000 (TrEMBLrel. 15, Created)
01-OCT-2000 (TrEMBLrel. 15, Last sequence update)
01-MAR-2001 (TrEMBLrel. 16, Last annotation update)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         Bacteria; Firmicutes; Bacillus/Clostridium group;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  317 AA
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       Bacillus/Staphylococcus group; Bacillus
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  PRT;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 SEQUENCE FROM N.A. STRAIN=ATCC 14579 TYPE STRAIN;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            PUTATIVE UV-ENDONUCLEASE.
   Conservative
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          566 --YDQT'RDG-----
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  PRELIMINARY;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         Bacillus cereus.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      NCBI_TaxID=1396
   241;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       SEQUENCE
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              Q9L4F0;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  Q9L4F0
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     RESULT 3
                                                                                           158
                                                                                                                                                                                                                                                                           276
    Matches
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     Q9L4F0
                                                                                                                                                                                  ДQ
                                                                                                                                                                                                                                            q
                                                                                                                                                                                                                                                                                                        q
                                                                                                                                                                                                                                                                                                                                                                  qq
                                                                                                                                                                                                                                                                                                                                                                                                  ŏ
                                                                                                                                                                                                                                                                                                                                                                                                                               qq
                                                                                                                                                                                                                                                                                                                                                                                                                                                            ŏ
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         qq
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        Qγ
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  õ
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  õ
                                                                                                                                                                                                                 ò
                                                                                                                                                                                                                                                                           δy
```

276 KPFLHIAKKHNHNFDIMIESKQKDLALFQL 305

a

```
8;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           STRAIN=ATCC 14579 TYPE STRAIN;
MEDLINE=2005637; Pubmed=10589720;
Okstad O., Gominet M., Purnelle B., Rose M., Lereclus D., Kolsto A.B.;
"Sequence analysis of three Bacillus cereus loci under PICR-regulated genes encoding degradative enzymes and enterotoxin.";
Microbiology 145:3129-3138(1999).
EMBL; AJ243712; CAB69814.1;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         291 QNVLDLIKLVEWNHNFGIHFMRVSSDLFPFASH--AKYGYTLEFAQSHLEEVGKLANKYN 348
                                                                               46 QNLEHTIRILHYNIAHEIPLYRLSSSIVPLATHPEVEFDYIGVFTPL-WRKIGALIKEHN 104
                                                                                                                          349 HRLTMHPGQYTQIASPREVVVDSAIRDLAYHDEILSRMKLNEQLNKDAVLIIHLGGTFEG 408
                                                                                                                                            PGTLREGSLDLMPLIPTIRETWTRKGITQKQHYSESADPTAISGMKRRAHSDRVFDFPPC 526
                                                                                                                                                                                                                                                                                221 EEPLEE-----LLPAIFETWSHTNISPKVHISSPR----SEKEFRAHAEYI-DLEFI 267
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             291 QNVLDLIKLVEWNHNFGIHFMRVSSDLFPFASH--AKYGYTLEFAQSHLEEVGKLANKYN 348
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  349 HRLTMHPGQYTQIASPREVVVDSAIRDLAYHDEILSRMKLNEQLNKDAVLIIHLGGTFEG 408
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            KKETLDRFRKNYQRLSDSVKARLVLENDDVSWSVQDLLPLCQELNIPLVLDWHHH--NIV 466
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   PGTLREGSLDLMPLIPTIRETWTRKGITQKQHYSESADPTAISGMKRRAHSDRVFDFPPC 526
                                 Gaps
                                                                                                                                                                                    KKETLDRFRKNYQRLSDSVKARLVLENDDVSWSVQDLLPLCQELNIPLVLDWHHH--NIV
                                32;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   11.5%; Score 359; DB 2; Length 325; ilarity 30.7%; Pred. No. 2.5e-15; Conservative 60; Mismatches 95; Indels
11.5%; Score 359; DB 2; Length 317; 30.7%; Pred. No. 2.4e-15;
                                Indels
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    325 AA; 37900 MW; 07EFBDD1FA580653 CRC64;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                      Last sequence update)
Last annotation update)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            Bacteria; Firmicutes; Bacillus/Clostridium group; Bacillus/Staphylococcus group; Bacillus.
NCBI_TaxID=1396;
                                                                                                                                                                                                                                                                                                                                                                                                                        325 AA
                              60; Mismatches
                                                                                                                                                                                                                                                                                                               DP-----TMDLMIEAKEKEQAVFEL 546
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                DP-----TMDLMIEÄKEKEQAVFEL 546
                                                                                                                                                                                                                                                                                                                             Created)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                01-OCT-2000 (TrEMBLrel. 15, 01-OCT-2000 (TrEMBLrel. 15, 01-MAR-2001 (TrEMBLrel. 16,
                              83; Conservative
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  PUTATIVE UV-ENDONUCLEASE
                                                                                                                                                                                                                                                                                                                                                                                                                        PRELIMINARY;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               Best Local Similarity
Matches 83; Conserv
                Best Local Similarity
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               [1]
SEQUENCE FROM N.A.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  Bacillus cereus.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       Endonuclease.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    SEQUENCE
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    Query Match
   Query Match
                                                                                                                                                                                                                                                                                                                                                                                                                      Q9L4E9
Q9L4E9;
                                                                                                                                                                                                                                                                                                                                                                                           4
                                                                                                                                                                                                                                                 467
                                                                                                                                                                                    409
                                                                                                                                                                                                                                                                                                             527
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   467
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  229
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                527
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      409
                                Matches
                                                                                                                                                                                                                                                                                                                                                                                           RESULT
                                                                                                                                                                                                                                                                                                                                                                                                           Q9L4E9
                                                                                                                                                                                                                                                                                                                                                                                                                                     g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           q
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 à
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  g
                                                                                                                                                      QQ
                                                                                                                                                                                                                                                                              a
                                                                                                                                                                                                                                                                                                               ŏ
                                                                                                                                                                                                                                                                                                                                         g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       ò
                                                              ò
                                                                                                                        ö
                                                                                                                                                                                    õ
                                                                                                                                                                                                                g
                                                                                                                                                                                                                                                 ò
```

```
302 WNHNFGIHFMRVSSDLFPFASH---AKYGYTLEFAQSHLEEVGKLANKYNHRLTMHPGQY 358
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        58 HNLAYQISFERLSSKLVPLVNHPLTEGWKYELAIAE-ELQAVGEFASEHQMRIDFHPDHF 116
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           359 TQTASPREVVVDSAIRDLAYHDEILSRMKLNEQLNKDAVLIIHLGGTFEGKKETLDRFRK 418
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            117 VVLNSEAKEITRRSLQTLLYHYKLLKGMEIDPRHR----CVLHVGGKKKGVEAGLEQFIE 172
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               419 NYQRLSDSVKARLVLENDDVSWSVQDLLPLCQELNIPLVLDWHHHNIVPGTLREGSLDLM 478
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 173 NTASIPKSLLSMIMLENDDKSYTIDDVLYLGEKLAIPVVLDIHHHDVL----HRSKSLQ 227
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 479 PLIPTIRETWTRKGITOKOHYSESADPTAISGMKRRAH-----SDRVFDF----PPCDP 528
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                228 ETWQRIVATWEDSPLPVKIHLS----SPLSGEDDPRHHDYINADRFIAFLHEIGADAVD 282
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     247 RGRLGYACLNTILRSMKERVFCSRTCRITTIQR-----DGLESVKQLGTQNVLDLIKLVE 301
                                                                                                                                                                                                                                                                                                                                        "Complete genome sequence of the alkaliphilic bacterium Bacillus halodurans and genomic sequence comparison with Bacillus subtilis."; Nucleic Acids Res. 28:437-4331(2000).
EMBL: AP001516; BAB06316.1: -.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      2 RIQFGYVAMSMELANASP----SKTMTATQFEKIEDHEAGLRKLERIAKTNLHNCLRLLK 57
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      Bacteria; Thermus/Deinococcus group; Deinococcales; Deinococcus.NCBL_TaxID=1299;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    37;
                                                                                                                                                                                                                                                                                          Takami H., Nakasone K., Takaki Y., Maeno G., Sasaki R., Masui
Fuji F., Hirama C., Nakamura Y., Ogasawara N., Kuhara S.,
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   Length 322;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    Indels
                                                                                                                                                                                                                                                                                                                                                                                                                Complete proteome.
SEQUENCE 322 AA; 36900 MW; 4B0DC9291CD1E7B4 CRC64;
                                                                                  Last sequence update)
Last annotation update)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         Last sequence update)
Last annotation update)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                              ; Score 306.5; DB 2;
; Pred. No. 5.3e-12;
75; Mismatches 123;
                                                                                                                                                              Bacteria; Firmicutes; Bacillus/Clostridium group;
Bacillus/Staphylococcus group; Bacillus.
NCBI_TaxID=86665;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         AA
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         305
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        Created)
                                                                  Created)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         PRT;
                                                                                                                                                                                                                                                                          MEDLINE=20512582; PubMed=11058132;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   529 IMDLMIEAKEKEQAVFELCR 548
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       : :|||||:|:|:| :
283 HLHVMIEAKKKDLALFQLMK 302
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  9.8%;
                                                            01-0CT-2000 (TrEMBLrel. 15, 01-0CT-2000 (TrimBLrel. 15, 01-0CT-2000 (TrEMBLrel. 15,
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        01-MAY-2000 (TrEMBLrel. 13,
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         01-MAY-2000 (TrEMBLrel. 13, 01-MAY-2000 (TrEMBLrel. 13,
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    Conservative
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         PRELIMINARY;
                                                                                                                                                                                                                                                    JCM 9153;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               Deinococcus radiodurans.
                                 PRELLIMINARY;
                                                                                                                                                      Bacillus halodurans.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  Best Local Similarity
                                                                                                                                                                                                                                         SEQUENCE FROM N.A.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           UV - ENDONUCLEASE.
                                                                                                                    BH2597 PROTEIN.
                                                                                                                                                                                                                                                                                                                          Horikoshi K.;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    85:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   Query Match
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        ,6M026Q
                                                Q9K9P8;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       6M0S60
                                Q9K9P8
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           UVSCDE
                                                                                                                                    BH2597
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       9
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    Matches
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       RESULT
RESULT
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       6M0S60
              Q9K9P8
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         ΩI
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 qq
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    QQ
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        qq
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            Dp
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               qq
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 ó
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   òχ
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    qq
```

6

```
Best Local Similarity
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                Q9ESK9
                                                                                                                                                                                                                                                                                                                                                                                                                                                  œ
                    Matches
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              Matches
                                                                                                                                                                                                                                                                                                                                                                                                                                                  RESULT
                                                                                                                                                                                                                                                                                                                                                                                                                                                                  Q9ESK9
                                                   δλ
                                                                                  9
                                                                                                                Qγ
                                                                                                                                                                            Óλ
                                                                                                                                                                                                          Ω
                                                                                                                                                                                                                                        δ
                                                                                                                                                                                                                                                                       q
                                                                                                                                                                                                                                                                                                      ΟŽ
                                                                                                                                                                                                                                                                                                                                   q
                                                                                                                                                                                                                                                                                                                                                                    ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                ACCOCCOS ON THE PROPERTY OF T
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            ōλ
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          qq
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         οy
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       Db
                                                                                                                                                                                                          10;
                              Kitayama S., Kikuchi M., Funayama T., Narumi I., Watanabe H.;
"Cloning of structural gene of an alternative incision enzyme for DNA
damage in Delnococcus radiodurans.";
Submitted (OCT-1999) to the EMBL/GenBank/DDBJ databases.
EMBL; AB0313747; BAA85759.1; -
                                                                                                                                                                                                                                                                                                                                                                                                                                                  249 RLGYACLNTILRSMKERVFCSRTCRITTIQRDGLESVKQLGTQNVLDLIKLVEWNHNFGI 308
                                                                                                                                                                                                                                                                                                   309 HFMRVSSDLFP---FASHAKYGYTLEFAQSHLEEVGKLANKYNHRLTMHPGQYTQIASPR 365
                                                                                                                                                                                                                                                                                                                                   68 RLYRLSSSLFPMLDLAGDDTGAAVLTHLAPQLLEAGHAFTDAGVRLLMHPEQFIVLNSDR 127
                                                                                                                                                                                                                                                                                                                                                                   366 EVVVDSAIRDLAYHDEI----LSRMKLNEQLNKDAVLIIHLGGTFEGKKETLDRFRKNY 420
                                                                                                                                                                                                                                                                                                                                                                                               128 PEVRESSVRAMSAHARVMDGLGLARTPWN-----LLLLHGGKGGRGAE-----LAALI 175
                                                                                                                                                                                                                                                                                                                                                                                                                               421 QRLSDSVKARLVLENDDVSWSVQDLLPLCQELNIPLVLDWHHHNIVPGTLREGSLDLMP- 479
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           480 -LIPTIRETWTRKGIT-----QKQHYSESADPTAISGMKRRAHSDRVFDFPPCDPTMD- 531
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 227 QEDPSVREWVLRARATWQPPEWQVVHLS----NGIEGPQDRRHSHLIADFPSAYADVPW 281
                                                                                                                                                                                                          Gaps
                                                                                                                                                                                                                                                                     10 OLGLVCLTVGPEVRFRTVTLSRYRALSPAEREA -- KLLDLYSSNIKTLRGAADYCAAHDI 67
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         White O., Eisen J.A., Heidelberg J.F., Hickey E.K., Peterson J.D., Dodson R.J., Haft D.H., Gwinn M.L., Nelson W.C., Richardson D.L., Moffat K.S., Qin H., Jiang L., Pamphile W., Crosby M., Shen M., Vamathevan J.J., Lam P., McDonald L., Utterback T., Zalewski C., Makarova K.S., Aravind L., Daly M.J., Minton K.W., Fleischmann K.E., Salzberg S., Smith H.O., Venter J.C.,
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 Bacteria; Thermus/Deinococcus group; Deinococcales; Deinococcus.
                                                                                                                                                                                                          45;
                                                                                                                                                                          1 7.4%; Score 229.5; DB 2; Length 305; Similarity 27.9%; Pred. No. 3.9e-07;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          "Genome sequence of the radioresistant bacterium Deinococcus
                                                                                                                                                                                                          42; Mismatches 140; Indels
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        C4EA0D0AD2C38988 CRC64;
                                                                                                                            305 AA; 33592 MW; B94D333243E2FEA4 CRC64;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 01-MAY-2000 (TrEMBLrel. 13, Created)
01-MAY-2000 (TrEMBLrel. 13, Last sequence update)
01-MAR-2001 (TrEMBLrel. 16, Last annotation update)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      326 AA.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      PRT;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              MEDLINE=20036896; Pubmed=10567266;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                DAMAGE ENDONUCLEASE, PUTATIVE.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  Endonuclease; Complete proteome
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      35693 MW;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        Science 286:1571-1577(1999).
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         EMBL; AE002022; AAF11370.1;
                                                                                                                                                                                                         88; Conservative
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         532 LMIEAKEKEOAVFEL 546
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      282 IEVEAKGKEEAIAAL 296
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                Deinococcus radiodurans.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     PRELIMINARY;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    326 AA;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              SEQUENCE FROM N.A.
SEQUENCE FROM N.A.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 NCBI_TaxID=1299;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    TIGR; DR1819; -
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         radiodurans R1.
                                                                                                              Endonuclease
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         Fraser C.M.;
                  STRAIN=KR1;
                                                                                                                             SEQUENCE
                                                                                                                                                                          Query Match
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        SEQUENCE
                                                                                                                                                                                           Local
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     O9RTE6
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     Q9RTE6
                                                                                                                                                                                                          Matches
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   RESULT
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      Q9RTE6
   RR RT RE DR KW KKW
                                                                                                                                                                                                                                       δ
                                                                                                                                                                                                                                                                       a
                                                                                                                                                                                                                                                                                                     δ
                                                                                                                                                                                                                                                                                                                                염
                                                                                                                                                                                                                                                                                                                                                                                                                                                            g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     Q
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        ōλ
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   ă
                                                                                                                                                                                                                                                                                                                                                                                               a
                                                                                                                                                                                                                                                                                                                                                                                                                             δ
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           δ
```

7.4%; Score 229.5; DB 2; Length 326;

Query Match

```
10;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        35;
                                Gaps
                                                                                                                                                                                                                                                                                                                                                                               248 QEDPSVREWVLRARATWQPPEWQVVHLS-----NGIEGPQDRRHSHLIADFPSAYADVPW 302
                                                                                        249 RLGYACLNTILRSMKERVFCSRTCRITTIQRDGLESVKQLGTQNVLDLIKLVEWNHNFGI 308
                                                                                                                                                                                                                     309 HFMRVSSDLFP---FASHAKYGYTLEFAQSHLEEVGKLANKYNHRLTMHPGQYTQIASPR 365
                                                                                                                                                                                                                                                                             89 RLYRLSSSLFPMLDLAGDDTGAAVLTHLAPQLLEAGHAFTDAGVRLLMHPEQFIVLNSDR 148
                                                                                                                                                                                                                                                                                                                                               36 EVVVDSAIRDLAYHDEI----LSRMKLNEQLNKDAVLIIHLGGTFEGKKETLDRFRKNY 420
                                                                                                                                                                                                                                                                                                                                                                                                                                                                     421 QRLSDSVKARLVLENDDVSWSVQDLLPLCQELNIPLVLDWHHHNIVPGTLREGSLDLMP- 479
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           480 -LIPTIRETWTRKGIT-----QKQHYSESADPTAISGMKRRAHSDRVFDFPPCDPFMD- 531
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        592 VQPFLRVPLLCDFEPLHQHVLALHNLVKAAQSLDEMSQTITDLLN---EQKVSTSQASPQ 648
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            64 MLPQVNIGA------NSFSAETPVDLKKENETELANISGPHKKSTSTSTRK 108
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      --- DKIDESVASYDSSTHLR 139
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 SAASPRIESTTG1TTTTSPKTPPPLTVQDTLCPAVCPLEELSPDSIDAHTFDFETISH-- 706
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            140 RSSRSKKPVNYNSSSESEEGISKATKKVKQKEEEETVVEEVDEKSLKNESSSDE--FEP 197
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    --PNTEQPVH-----QASIDLDSLAESPENSAVNEFVIEENLSSPNPISDP 753
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              532 SKRESFGKLFRKSFLRNRLFKGLDSWPSSFCTQKPRKFDCELPDISLKDLQFLQSFCPSE 591
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           Gaps
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                12 SKRIVF-----TILKQKAFKGNHPCVPSVCTITYSRFHC-LPD-TLKSLLPMSSKTTLS 63
                                                                                                                                                   31 QLGLVCLTVGPEVRFRTVTLSRYRALSPAEREA - - KLLDLYSSNIKTLRGAADYCAAHDI
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
                                45;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        255; Indels 220;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         DB 11; Length 1588;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     Chano T., Okabe H., Saeki Y.; "Nuclear protein CC1, mouse homologue of human KIAA0203."; Submitted (OCT-2000) to the EMBL/GenBank/DDBJ databases.
                                42; Mismatches 140; Indels
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             182276 MW; 930265D58A21DF0F CRC64;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               01-MAR-2001 (TrEMBLrel. 16, Created)
01-MAR-2001 (TrEMBLrel. 16, Last sequence update)
01-JUN-2001 (TrEMBLrel. 17, Last annotation update)
27.9%; Pred. No. 4.2e-07;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      5.3%; Score 164; DB 1.
20.0%; Pred. No. 0.045;
tive 113; Mismatches 29
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       PRT; 1588 AA.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 MGD; MGI:1341850; Ccl.
InterPro; IPR000626; Ubiquitin.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        Conservative 113;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    EMBL; AB050017; BAB16846.1;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   109 RARSSKKKATDSVS-----
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   Pfam; PF00240; ubiquitin;
                                88; Conservative
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     532 LMIEAKEKEOAVFEL 546
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         ::||| ||:||: |
303 IEVEAKGKEEAIAAL 317
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          PRELIMINARY;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              TISSUE=SKELETAL MUSCLE;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      Mus musculus (Mouse).
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           COILED-COIL PROTEIN.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             1588 AA;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       Similarity
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             SEQUENCE FROM N.A.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              NCBI_TaxID=10090;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 147;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 SEQUENCE
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         Query Match
Best Local S
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        OBESK9;
```

. 8

198 VVPEQLETPISKRRRSRSSAKNLEKESTMNLDDHAPREM 236	237 FDCLDKPIPWRGRLGYACLNTILRSMKERVFCSRTCRITTIQRDGLESVKQLG 289 1 :	290 TONVLD-LIKLVEWNHNFGIHFMRVSSDLFPFASHARYGYTLEFAQSHLEEVGKLANKYN 348 :	349 HRLTMHPGQYTQIASPREVVV	377 AYHDEILSRWKLNEQLNKDAVLIIHLGGTFEGKKFILDRFRK-NYQRLS 424	425 DSVKARLVLENDDVSWSVQDLLPLCQELNIPLVLDWHHHNIVPGTLREGSLDLM 478 ::	479 PLIPTIRETWTRKGITQKQHYSESADPTAISGMKRRAHSDRVFDFPPCDPTMDLMIEA 536 1071 KILLEESRTQQKEMLKSLLEQETENLRTEISKLNQKIHDNNESYQVGLSELRALMT 1126	537 KEKEQAVFELCRRYELONPPCPLEIMGPEYDQTRDGYYPPGAEKRLTARKRR 588	589 SRKEEVEEDEK 599 :: 1179 LNSELSALEKQKDEK 1193	045198 PRELIMINARY; PRT; 368 AA. 045198 PRELIMINARY; PRT; 368 AA. 045198. 01-JUN-1998 (TrEMBLrel. 06, Last sequence update) 01-JUN-1998 (TrEMBLrel. 06, Last sequence update) 01-JUN-1998 (TrEMBLrel. 06, Last sequence update) 01-NOV-1998 (TrEMBLrel. 06, Last sequence Rhabditidas; Rooper 10, Cooper 10, Cooper 10, Coulson A., STREINE-811507 B., Durbin R., Pavello A., Editor M., Johnston L., Gardner A., Green P., Hawkins T., Hillier L., Jier M., Johnston L., Gardner A., Green P., Hawkins T., Hillier N., Laister N., Sanchammer E., Staden R., Salston J., Snownkeen R., Thierry-Mieg J., Thomas K., Vaudin M., Vaughan K., Waterston R., Wallston A., Wallston A., Wallston S., Staden R., Salston J., Wohldman P., Rature 368:32-38(1994). 121 121 122 122 123
oy	Qy	Qy	Qy	Qy	Qy	Qy	Oy	Oy	RESULTA 10 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0
B	Db	Db	Db	Db	Db	Db	Db	Db	

```
29;
                                                                                                                                                                                                                                73 NSFSAETPVDLKK-----ENETELANISGPHKKSTSTSTRKRARSSKKKATDSVSDK 124
                                                                                                                                                                                                                                                      145 SDSSDSSDSSDDDKKKKKKKKKKKKKKKKKKKONKGSSDSSDDSSDDEKPKKKKKKKKKKKAPK 204
                                                                                                                                                                                                                                                                                                                                                                           SEEQIS-KATKKVKQKEE------196
                                                                                                                                                                                                                                                                                                                                                                                                 1:1:: | || |:| | |-| 205 SDEEVTIKIIKKPKKKPEDLERRRHRVKKTEEHEEEKD-KEEHHESNDENMEFDTRARGA 263
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        |: :| | |: |: | | |: |
841 PVVTKAALKAVGRPKRTPGRRR------KKDRGASPSPSPSPSPPPPFTTLTPIVLPR 891
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         108 KRARSSKKKAT------DSVSDKIDESVASYDSSTHLRRSSRSKKPVNY 150
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             151 NSSSESESEEQISKATKKVKQKEEEEYVEEVDEKS------LKNESSSDEF 195
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  55 PMSSKTTLSML--PQVNIGANSFSAETPVDLKKENETELANISGPHKKSTSTST----R 107
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            Query Match 5.0%; Score 155; DB 5; Length 1641; Best Local Similarity 19.2%; Pred. No. 0.17; Matches 126; Conservative 91; Mismatches 227; Indels 212; Gaps
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     Eukaryota; Metazoa; Nematoda; Chromadorea; Rhabditida; Rhabditoidea;
Rhabditidae; Peloderinae; Caenorhabditis.
                                                                                                                                                             DB 5; Length 368;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   none; "Genome sequence of the nematode C.elegans: A platform for
                                                                                                                                                                                                Indels
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      EMBL; AL132898; CAC14406.1; -.
InterPro; IPR002965; P_rich_extensn.
PRINTS; PR01217; PRICHEXTENEM; F6C38533AA158DID CRC64;
SEQUENCE 1641 AA; 181927 MW; F6C38533AA158DID CRC64;
                                                   Waterston R.;
Submitted (FEB-1998) to the EMBL/GenBank/DDBJ databases.
EMBL: AF007663; AAC04447.1; -
SEOUENCE 368 Aa; 41823 MW; AE74C475F8340CDA CRC64;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          Subston J.E.; Submitted (OCT-2000) to the EMBL/GenBank/DDBJ databases.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         01-WAR-2001 (TrEMBLrel. 16, Created)
01-WAR-2001 (TrEMBLrel. 16, Last sequence update)
01-JUN-2001 (TrEMBLrel. 17, Last annotation update)
                                                                                                                                                           Query Match 5.0%; Score 156.5; DB 5; Best Local Similarity 26.1%; Pred. No. 0.021; Matches 57; Conservative 37; Mismatches 61;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                    197 PVVPEQLETPISKR-RRSRSSAKNLEKESTMNLDDHAP 233
                                                                                                                                                                                                                                                                                                      125 IDESVASYDSSTHLRRSSRSKKPVNYNSSSES----
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             [2]
SEQUENCE FROM N.A.
MEDLINE-99069613; Pubmed-9851916;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   investigating biology.";
Science 282:2012-2018(1998).
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          PRELIMINARY;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  Caenorhabditis elegans.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           SEQUENCE FROM N.A.
                SEQUENCE FROM N.A. STRAIN=BRISTOL N2;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 Y59A8B.1 PROTEIN.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          NCBI_TaxID=6239;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              Query Match
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  Y59A8B.1
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            09GRZ3;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          Q9GRZ3
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        RESULT 10
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     264
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         Q9GRZ3
RN
RC
RA
RL
DR
SQ
                                                                                                                                                                                                                                  ŏ
                                                                                                                                                                                                                                                                   q
                                                                                                                                                                                                                                                                                                        δy
                                                                                                                                                                                                                                                                                                                                        q
                                                                                                                                                                                                                                                                                                                                                                             õ
                                                                                                                                                                                                                                                                                                                                                                                                          qq
                                                                                                                                                                                                                                                                                                                                                                                                                                               ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 q
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            ŏ
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    QQ
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         Qy
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          Пр
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             ò
```

```
37;
                                                                                     1005 IFVEPSTLEKPERRNGTTEERVMTADQRAMFDAKIEEARRSRMSTRDCSVVSTLGPVKSK 1064
                                                                                                                                                                 1185 KTLSHPDYVRSKNEAEKAEFHGK-GGTMRITNRNLKMLTRQFDLPKMSSRFRKFVRIRRH 1243
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     1302 DQFEWLAKNSPNTQVKVGSLTNKQFIETMPMKKYYESAMETLDMGTFRFGPLMS--LSMV 1359
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     125 HHLARKNELKEBANVARRSEILKAVVKROREIGVPTTLMSKSAIELVEEDEKERKNHKNN 1184
                                                                                                                                                                                                                                                                                                                                                                                                                      1244 PNGMATIISCDYNQIKQHLGPNEMKHFERQFVRLGFAENNGVPLFAIGVMENAAE--ALH 1301
                                                           ----ISKRRRSRSSAKNLEKESTM----- 226
                                                                                                                                                                                                                      257 TILRSMKERVFCSRTCRITTIQRDGLESVKQLGTQNVL---DLIKLVE-----wnHNFG 307
                                                                                                                                                                                                                                                                                                   308 ----IHFMRVSSDLFPFASHAKYGYTLEFAQSHLE-----EVGKLANKYNH--RLTMH 354
                                                                                                                                                                                                                                                                                                                                                                                  355 PG------QYTQI---ASPREV-------------VVDSAIRDLAYH 379
                                                                                                                                                                                                                                                                                                                                                                                                                                                              380 DEI-----LSRMKLNEQLNKDAVLIIHLGGTFEGKKETLDRFRKNYQRLSDSVKARL 431
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             432 VLENDDVSWSVQDLLPLCQELN----IPLVLDWHHHNIVPGTLREGSLDLMPLIPTIRET 487
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      ---KKPVNYNSSSESESEQISKATKKVKQKEEEEYVEEVDEKSLKNE- 189
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     228 NINILKNEGQLVAQKEEFDYDENMEKAKEEK-QKALEKIGKETDEEPFMFTDDKFLENQV 286
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 -----SSSDEFEPVVP----EQLE-TPISKRRSRSSAKNLEKESTM-----NLDDHA 232
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  43 FHCLPDTLKSLLPMSSKTTLSMLPQVNIGANSFSAETPVDLKKENETELANISGPHKKST 102
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      122 FKASPSVVKTSTPTGTHTS-----GSKSSS---PSSTKSSSPSNVKTAS-PHGESN 168
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               103 STSTRKRARSSKKKATDSVSDKIDESVAS------YDSSTHLRRSSRS---- 144
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          94; Mismatches 240; Indels 221; Gaps
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               488 WTRKG----ITQKQHYSESADPTAISGMKRR-----AHSDRVFDFFPCDPT 529
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               5.0%; Score 154.5; DB 5; Length 782; 20.9%; Pred. No. 0.074;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               Alveolata; Apicomplexa; Haemosporida; Plasmodium.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                Lal A.A.;
ed (FEB-1995) to the EMBL/GenBank/DDBJ databases.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       Submitted (FEB-1995) to ....
EMBL; U20986; AAA53682.1; -.
FORCENTE 782 AA; 90160 MW; DE6D1BE2FAC308A9 CRC64;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      NOV-1996 (TrEMBLrel. 01, Created)
-NOV-1996 (TrEMBLrel. 01, Last sequence update)
-JUN-2000 (TrEMBLrel. 14, Last annotation update)
                                                                                                                                        227 ---NLDD--HAPREMFDCLDKPIPWRGRL------
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    782 AA.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    PRT;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    RHOPTRY ASSOCIATED PROTEIN 1.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   Query Match 5.0%;
Best Local Similarity 20.9%;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          Matches 147; Conservative
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    PRELIMINARY;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            lasmodium reichenowi.
                                                           196 EPVVPEQLETP----
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           SEQUENCE FROM N.A.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             Eukaryota; Alveos
NCBI_TaxID=5854;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           01-NOV-1996
01-JUN-2000
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    026104
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          190
                                                                                                                                                                                                                                                                                                                                     qq
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   qq
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               qq
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   pp
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           δλ
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  g
                                                                                                                                          ă
                                                                                                                                                                                                                    δ
                                                                                                                                                                                                                                                        g
                                                                                                                                                                                                                                                                                                   δλ
                                                                                                                                                                                                                                                                                                                                                                                  ŏ
                                                                                                                                                                                                                                                                                                                                                                                                                      qq
                                                                                                                                                                                                                                                                                                                                                                                                                                                            ōλ
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            ŏ
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    δ
```

```
39;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   94; Mismatches 241; Indels 219; Gaps
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           233 PREMFDCLDKPIPWRGRLGYACLNTILRSMKERV--FCSRTCRITTIQRDGLESVKQLGT 290
                                                                                                                                                              291 QNVLDL-IKLVEWNH-----NFGIHFMRV---SSDLFPFASHAKY-----GYTLE----F 332
                                                                                                                                                                                                        396 -- VLDLELKDTECKFGTCIGSFGEHHLRLYEFENDLFKFHPNIDYLTLADGYKLQKNHIY 453
                                                                                                                                                                                                                                                                                                      454 ELSHVNFCLLNPKTLEEFLKKKEIKDLMGGDDLIKYKENFDNFMSISLTCHIESLIYDDI 513
                                                                                                                                                                                                                                                                                                                                                                                              514 EASQDVAAVLKIAKSKLHVITSGLSYKARKLVYKIYSEIQKNPDELYEKLTWIYDNIYML 573
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  S---ADPTAIS----GMKRRAHSDRVFDFPPCDPTMDLMIEAKEKEQAVFELCRRYELQN 554
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 669 SHVYYDPTVASYAYYNLDRRTMVTIINDY-----FEAKKKELTVIVSRMKTDM-- 716
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     43 FHCLPDTLKSLLPMSSKTTLSMLPQVNIGANSFSAETPVDLKKENETELANISGPHKKST 102
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              103 STSTRKRARSSKKKATDSVSDKIDESV--ASYDSSTHLRRSSRSKKPVNYN----SSSES 156
                                                                                                                 347 ETMLYNCPNN-----SHLFDTISSLQGRLIDIKKRENMISTTFEQQKECLKNMG- 395
                                                                                                                                                                                                                                                         333 AQSH-----LEEVGKLAN-----KYNHR------LTMHPGQ--YTQI 361
                                                                                                                                                                                                                                                                                                                                                  362 ASPREV-----VVDSAI----RDLAYH-----DEILSRMKL---NEQLN 393
                                                                                                                                                                                                                                                                                                                                                                                                                                             394 KDAVLIIHLGGT---FEGKKETLDRFRKNYQRLSDSVK-----ARLVLEN---DDVSWS 441
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       574 KRYYTAYALEGUCSYIENDKTQMDTELHIYNKIVDSVRYYSSCFKNVIVYNAIMSGIHEK 633
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      442 VQDLLPLCQELNIPLVLDWHHHNIVPGTLREGSLDLMPLIPTIRETWTRKGITQKQHYSE 501
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   634 IKHFLKLVPRQN--FLLDYHFNSIF-------EKEIKPVKKYST 668
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               12 FKASPSVVKTSTPSGTQT-----SGLKSSSPSSTKSSSPSNVKSAS-PHGESN 168
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       157 ESEEQISKATKK---VKQKEEEEYVEEVD-----EKSLKNESSSDEFEPVVPEQ---LE 204
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         4.8%; Score 149.5; DB 5; Length 782; 21.0%; Pred. No. 0.15;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                Alveolata; Apicomplexa; Haemosporida; Plasmodium.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               Li X.R., Luo S.H., Yu X.B., Shan Z.X., Ma C.L.;
Submitted (NOV-1999) to the EMBL/GenBank/DDBJ databases.
EMBL; AF206631; AAF15365.1;
SEQUENCE 782 AA: 90082 MW; BE1F4CF2883903FD CRC64;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     01-MAY-2000 (TrEMBLrel. 13, Created)
01-MAY-2000 (TrEMBLrel. 13, Last sequence update)
01-MAY-2000 (TrEMBLrel. 13, Last annotation update)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             555 PPCPLEIMGPEYDQTRDGYYPPGAEKRLTARKRRSRKEEVEE 596
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          717 ----LSLQNDESKIPND----KSANSKLATRLMRKFKAEIRD 750
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           PRT;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             RHOPTRY ASSOCIATED PROTEIN-1.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        Matches 147; Conservative
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           PRELIMINARY;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          Plasmodium falciparum.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  Best Local Similarity
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     SEQUENCE FROM N.A.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                Eukaryota; Alveol
NCBI_TaxID=5833;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              STRAIN-FCC1/HN;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            Query Match
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       090414;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           090414
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            RESULT 12
Q9U414
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             169
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       Pp
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   q_{\Box}
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             ΟŊ
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       pp
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                ÓΫ
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        Ωp
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         δλ
                                                                     δ
                                                                                                               q
                                                                                                                                                                δλ
                                                                                                                                                                                                          qq
                                                                                                                                                                                                                                                       ò
                                                                                                                                                                                                                                                                                                   g
                                                                                                                                                                                                                                                                                                                                                  δ
                                                                                                                                                                                                                                                                                                                                                                                            q
                                                                                                                                                                                                                                                                                                                                                                                                                                             ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      Óγ
```

```
TPISKRRRSRSSAKNLEKESTMNLDDHAPREM------FDCLDKPIPWRG----- 248
                         284 NOVKERNVAGSFSRFFSKLNPFKKDEVIEKTEVSKKTFSGIGFNLTDKEAKVLGVGVTYQ 343
                                                      ----RLGYACLNT----ILRSMKERV--FCSRTCRITTIQRDGLESVKQLGTQNVLDLI 297
                                                                         344 EYPETMLYNCPNNSNLFDTIESLQGRIIDIKKRESMISTTFEQQKECLKNMG---VLDL- 399
                                                                                                                                            456
                                                                                                                                                                   H-----LEEVGKLAN-----KYNHR-----LTMHPGQ--YTQIASP 364
                                                                                                                                                                                               457 HVNFCLLNPKTLEEFLKKKEIKDLMGGDDLIKYKENFDNFMSISITCHIESLIYDDIEAS 516
                                                                                                                                                                                                                        365 REV-----EQLN--KDAVLII- 400
                                                                                                                                                                                                                                                                               401 --HLGGTFEGKKETLDRFRKN-----YQRLSDSVK-----ARLVLENDDVSW---SV 442
                                                                                                                                                                                                                                                                                                        575 RYYTAYALEGVCSYLEHDKSQMHTELHIYNKIVDSVRYYSSCFKNVIVYNAIISGIHEKI 634
                                                                                                                                                                                                                                                                                                                                            502
                                                                                                                                                                                                                                                                                                                                                                                               503 ---ADPTAIS----GMKRRAHSDRVFDFPPCDPTMDLMIEAKEKEQAVFELCRRYELQNP 555
                                                                                                                                                                                                                                                                                                                                                                                                                          670 HIYFDPTVASYAYYNLDRRTMVTIINDY------FEAKKKELTVIVSRMKTDM--- 716
                                                                                                             298 KLVEWNH-----GYTLE----FAQS
                                                                                                                                       400 ---ELNDTQCKFGTCIGSFGEHHLRLYEFENDLFKFHPNIDYLTLADGYKLQKNHIYELS
                                                                                                                                                                                                                                                                                                                                      443 QDLLPLCQELNIPLVLDWHHHNIVPGTLREGSLDLMPLIPTIRETWTRKGITQKQHYSES
                                                                                                                                                                                                                                                                                                                                                                  635 KHFLKLVPRHN--FLLDYHFNSIF-------EKEIKPAKKYSTS
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 Saccharomyces cerevisiae (Baker's yeast).
Eukaryota; Fungi; Ascomycota; Saccharomycotina; Saccharomycetes;
Saccharomycetales; Saccharomycetaceae; Saccharomyces.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            Pauley A.;
Submitted (NOV-1994) to the EMBL/GenBank/DDBJ databases.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             01-NOV-1996 (TrEMBLrel. 01, Created)
01-NOV-1996 (TrEMBLrel. 01, Last sequence update)
01-JUN-2001 (TrEMBLrel. 17, Last annotation update)
SIMILARITY TO ROD-LIKE TAIL OF SEVERAL MYOSINS.
IMH1 OR L2142.5 OR YLR309C.
                                                                                                                                                                                                                                                                                                                                                                                                                                                    556 PCPLEIMGPEYDQTRDGYYPPGAEKRLTARKRRSRKEEVEE 596
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                717 ---LSIONEESKIPND----KSANSKLATRLMKKFKAEIRD 750
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         PRT;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      PRELIMINARY;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                SEQUENCE FROM N.A.
STRAIN-S288C (AB972);
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       SEQUENCE FROM N.A.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       Nature 387:0-0(0).
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            NCBI_TaxID=4932;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  006704;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    206704
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          RESULT 13
                                                                                                                                                                   336
                          a
                                                                                                           ò
                                                                                                                                       a
                                                                                                                                                                   ò
                                                                                                                                                                                            g
                                                                                                                                                                                                                                                   a
                                                                                                                                                                                                                                                                                                          QQ
                                                                                                                                                                                                                                                                                                                                                                                                                        g
                                                      ò
                                                                               g
                                                                                                                                                                                                                        δ
                                                                                                                                                                                                                                                                               ö
                                                                                                                                                                                                                                                                                                                                       ò
                                                                                                                                                                                                                                                                                                                                                                  요
                                                                                                                                                                                                                                                                                                                                                                                            ó
                                                                                                                                                                                                                                                                                                                                                                                                                                                    ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                             Db
```

```
21;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         202 KEEVSISE-----LKEELAIKNHSLEDSRMKITELEQN------LSSKSTIMEE 244
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     1 | 1::|| | 1::|| | 1 | 1:: | 1: | 1:: | 1:: | 1:: | 1:: | 1:: | 1:: | 1:: | 1:: | 1:: | 1:: | 1:: | 1:: | 1:: | 1:: | 1:: | 1:: | 1:: | 1:: | 1:: | 1:: | 1:: | 1:: | 1:: | 1:: | 1:: | 1:: | 1:: | 1:: | 1:: | 1:: | 1:: | 1:: | 1:: | 1:: | 1:: | 1:: | 1:: | 1:: | 1:: | 1:: | 1:: | 1:: | 1:: | 1:: | 1:: | 1:: | 1:: | 1:: | 1:: | 1:: | 1:: | 1:: | 1:: | 1:: | 1:: | 1:: | 1:: | 1:: | 1:: | 1:: | 1:: | 1:: | 1:: | 1:: | 1:: | 1:: | 1:: | 1:: | 1:: | 1:: | 1:: | 1:: | 1:: | 1:: | 1:: | 1:: | 1:: | 1:: | 1:: | 1:: | 1:: | 1:: | 1:: | 1:: | 1:: | 1:: | 1:: | 1:: | 1:: | 1:: | 1:: | 1:: | 1:: | 1:: | 1:: | 1:: | 1:: | 1:: | 1:: | 1:: | 1:: | 1:: | 1:: | 1:: | 1:: | 1:: | 1:: | 1:: | 1:: | 1:: | 1:: | 1:: | 1:: | 1:: | 1:: | 1:: | 1:: | 1:: | 1:: | 1:: | 1:: | 1:: | 1:: | 1:: | 1:: | 1:: | 1:: | 1:: | 1:: | 1:: | 1:: | 1:: | 1:: | 1:: | 1:: | 1:: | 1:: | 1:: | 1:: | 1:: | 1:: | 1:: | 1:: | 1:: | 1:: | 1:: | 1:: | 1:: | 1:: | 1:: | 1:: | 1:: | 1:: | 1:: | 1:: | 1:: | 1:: | 1:: | 1:: | 1:: | 1:: | 1:: | 1:: | 1:: | 1:: | 1:: | 1:: | 1:: | 1:: | 1:: | 1:: | 1:: | 1:: | 1:: | 1:: | 1:: | 1:: | 1:: | 1:: | 1:: | 1:: | 1:: | 1:: | 1:: | 1:: | 1:: | 1:: | 1:: | 1:: | 1:: | 1:: | 1:: | 1:: | 1:: | 1:: | 1:: | 1:: | 1:: | 1:: | 1:: | 1:: | 1:: | 1:: | 1:: | 1:: | 1:: | 1:: | 1:: | 1:: | 1:: | 1:: | 1:: | 1:: | 1:: | 1:: | 1:: | 1:: | 1:: | 1:: | 1:: | 1:: | 1:: | 1:: | 1:: | 1:: | 1:: | 1:: | 1:: | 1:: | 1:: | 1:: | 1:: | 1:: | 1:: | 1:: | 1:: | 1:: | 1:: | 1:: | 1:: | 1:: | 1:: | 1:: | 1:: | 1:: | 1:: | 1:: | 1:: | 1:: | 1:: | 1:: | 1:: | 1:: | 1:: | 1:: | 1:: | 1:: | 1:: | 1:: | 1:: | 1:: | 1:: | 1:: | 1:: | 1:: | 1:: | 1:: | 1:: | 1:: | 1:: | 1:: | 1:: | 1:: | 1:: | 1:: | 1:: | 1:: | 1:: | 1:: | 1:: | 1:: | 1:: | 1:: | 1:: | 1:: | 1:: | 1:: | 1:: | 1:: | 1:: | 1:: | 1:: | 1:: | 1:: | 1:: | 1:: | 1:: | 1:: | 1:: | 1:: | 1:: | 1:: | 1:: | 1:: | 1:: | 1:: | 1:: | 1:: | 1:: | 1:: | 1:: | 1:: | 1:: | 1:: | 1:: | 1:: | 1:: | 1:: | 1:: | 1:: | 1:: | 1:: | 1:: | 1:: | 1:: | 1:: | 1:: | 1:: | 1:: | 1:: | 1:: | 1:: | 1:: | 1:: | 1:: | 1
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             112 SSKKK-----ATDSVSDK--IDESVASYDSSTHLRRSSRSKKPVNYNSSSESEE-EQIS 163
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                293 NKGKKNKGGITTGDISEEETVDNSINTEE------YDKLKENLQELQEKY 336
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   445 YEAKNTELRSKIE------LLSKKVEHLKNLCTEKEKEGTTSQNKVAKLNEEIS 492
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          63 --SMLPQVNIGANSFSAETPVDLKKENETELANISG------PHKKSTSTSTRKRAR 111
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   164 KATKKVKQKEEEEYVEEVDEKSLKN---ESSSDEFEPVVPEQLETPISKRRS----- 213
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     397 DMLRTVGNELVDAKDEIKESSSKQNEEVKTVKLELDDLRHKNATMIE-------A 444
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       305 NFGIHFMRVSSDLFPFASHAK------YGYTLEFAQSHLEEVGKLANKYNH 349
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            -----DIAKNSNSYEEYLKENGKLSERLN----- 590
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           591 ---ILQEKYNTLQNVKSNSNEHIDSIKRQCEELNVKLKESTKKILSLEDELN----E 640
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    252 YACLNTILRSMKERVFCSRTCRITTIQRDGLESVKQLGTQNVLD------LIKLVEWNH 304
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         RL-----TMHPGQYTQIASPREVVVDSAIRDLAYHDEILSRMKLNEQLNKDAVLIIH 401
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    402 LGGTFEGKKETLDRFRKNYQRLSDSVKARLVLENDDVSWSVQDLLPLCQELNIPLVLDWH 461
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   Gaps
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        6 KRNIQISKRIVFTILKQKAFKGNHPCVPSVCTITYSRFHCLPDTLKSLLPMSSKTTL--- 62
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               Eukaryota; Metazoa; Nematoda; Chromadorea; Rhabditida; Rhabditoidea;
Rhabditidae; Peloderinae; Caenorhabditis.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      493 OLTYEKSNITKELTSLRTSYKQKEKTVSYLEEQVKOFSEOKDVAEKSTEQLRKDHAKISN
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                Indels 185;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             Length 911;
                                                                                           Submitted (NOV-1994) to the EMBL/GenBank/DDBJ databases
                                                                                                                                                                                                                                                                    to the EMBL/GenBank/DDBJ databases
                                                                                                                                                                                                                                                                                                                                                                                                                                      090BD9AA5702C01D CRC64;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             Last sequence update)
Last annotation update)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      4.8%; Score 148.5; DB 3;
20.4%; Pred. No. 0.21;
ive 75; Mismatches 180;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      PRT; 1359 AA.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             Created)
                                                                                                                                                                                                                                                                                                                                                                                                                                      105224 MW;
                                                                                                                                                                                                                                                          Submitted (AUG-1997) to the EMBL: U1747; AA657359.1; -. SGD; S0004300; IMH1. InterPro: JPR000237; GRIP. Pfam; PF01465; GRIP; 1.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     20.4%;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      04,
08,
17,
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     Matches 113; Conservative
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   PRELIMINARY;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      01-JUL-1997 (TrEMBLrel. 01-NOV-1998 (TrEMBLrel.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         01-NOV-1998 (TrEMBLrel.
01-JUN-2001 (TrEMBLrel.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        Caenorhabditis elegans.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     : ||| || ;
641 YANIVQDKTREAN 653
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              462 HHNIVPGTLREGS 474
                                                                                                                                                           SEQUENCE FROM N.A.
STRAIN=S288C (AB972);
                                   STRAIN=S288C (AB972);
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        553 RLDLLKKENETLHN--
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     Best Local Similarity
SEQUENCE FROM N.A
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     B0041.7 PROTEIN.
                                                                  Waterston R.
                                                                                                                                                                                                                                     Cherry J.M.;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  Query Match
                                                                                                                                                                                                                                                                                                                                                                                                                                   SEQUENCE
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          002061;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   002061
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                RESULT 14
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       002061
ŏ
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        q
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             QΩ
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            qq
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               δy
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               qq
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     δ
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     qq
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        ŏ
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    qq
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      qd
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 QQ
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       qq
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   AC DT DT OCC OCC OCC OCC
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        ŏ
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          δ
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      qq
```

œ

```
224 STMNLDDHAPRE 235
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        186 LKNESSSD----
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  SEQUENCE
                                                                                                                                                                                                                                                                                                                                                                                                                     09U7E0;
                                                                                                                                                                                                                                                                                                                                                                                                09U7E0
                                                                                                                                                                                                                                                                                                                                                                    RESULT 15
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              Matches
                                                                                                                                                                                                                                                                                                                                                                                     09U7E0
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             ζý
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           a
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        οy
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        q
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        δ
                                                                                                                                                                                      qq
                                                                                                                                                                                                                         δλ
                                                                                                                                                                                                                                                                                      Óλ
                                                                                                                                                                                                                                                                                                                                                                                                                                  q
                                                                                           Qy
                                                                                                                         qq
Ω
                              ΩŽ
                                                            Wilson R., Aliascough R., Anderson K., Baynes C., Berks M.,.
Wilson R., Aliascough R., Anderson K., Eavelo J., Cooper J., Coulson A.,
Craxton M., Dear S., Du Z., Durbin R., Favello A., Fulton L.,
Gardner A., Green P., Hawkins T., Hillier L., Jier M., Johnston L.,
Jones M., Kershaw J., Kirsten J., Laister N., Latreille P.,
Lightning J., Lloyd C., Mcmurray A., Mortimore B., O'Callaghan M.,
Parsons J., Percy C., Rifken L., Ropera A., Saunders D., Shownkeen R.,
Smaldon N., Smith A., Sonnhammer E., Staden R., Sulston J.,
Thierry-Mieg J., Thomas K., Vaudin M., Vaughan K., Waterston R.,
Watson A., Weinstock L., Wilkinson-Sproat J., Wohldman P.,
"2.2 Mb of contiguous nucleotide sequence from chromosome III of C.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      Ouery Match 4.8%; Score 148.5; DB 5; Length 1359;
Best Local Similarity 20.9%; Pred. No. 0.36;
Matches 136; Conservative 96; Mismatches 222; Indels 197; Gaps
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       229 VKKESESE---DEAPEKK--KTEKRKRSKTSSEESSE--SEKSDEEEEEKSSP----KP 277
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              EMFDCLDKPIPWRGRLGYACLNTI-----LRSMKERVFCSRTCRITTIQRDGL----E 283
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          397 ERRKRLEKK------QKEFNGIVLEEGEDLTEMLTGTSSQRKLKSVVLDPDSSTVDEE 448
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           SVKQLGTQNVLDLIKLVEWNHNFGIHFMRVSSDLFPFASHAKYGYTLEFAQ----- 334
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         449 SKKPVEVHN--SLVRILKPHQAHGIQFM------YDCAFESLDRLDTEGS 490
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         335 ----SHLEEVGKLAN--KYNHRLTMHP--GQYTQ---IASPREVVVD---SAIRDLAYHD 380
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    83 LKKENETELANISGPHKKSTSTSTRKRARSSKKKATDSVSDKIDESVASYDSSTHLRRSS 142
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  143 RSKKPVNYNS-SSESEEQ-------ISKATKKVKQKEEEEYVEEVDEKS 185
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   278 KKKKPLAVKKLSSDEESEESDVEVLPQKKKRGAVTLISDSEDEKDQKSESE-ASDVEEKV 336
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 186 LKNESSSDEFEPVVPEQLETPISKRRRSRSSAKNLEKESTMNLD-------DHAPR 234
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              337 SKKKAKKQESSESGSDSSEGSITVNRKSKKKEKPEKKKKGIIMDSSKLQKETIDAERAEK 396
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         ATP-binding; Helicase.
SEQUENCE 1359 AA; 156235 MW; 5F2ACE099EDEDBF4 CRC64;
                                                                                                                                                                                                                                                                                                                                                      "The sequence of C. elegans cosmid B0041."; Submitted (JUN-1997) to the EMBL/GenBank/DDBJ databases.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  Waterston R.;
Submitted (JUN-1998) to the EMBL/GenBank/DDBJ databases.
-!- SIMILARITY: TO HELICASE C-TERMINAL DOMAIN.
                                                                                                                                                                                                                                                                                                                                                                                                                                      Waterston R.;
Submitted (APR-1997) to the EMBL/GenBank/DDBJ databases
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          ::<u>:</u>:
                                                  STRAIN=BRISTOL N2;
MEDLINE=94150718; PubMed=7906398;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           Interpro: IPR001410; DEAD.
Interpro: IPR001650; Helicase_C.
Interpro: IPR001630; SNF2_N.
Pfam; PF00271; helicase_C; 1.
Pfam; PF00176; SNF2_N; 1.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            <del>-</del>
..
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               EMBL; AF000196; AAC24256.1; -.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         SMART; SM00487; DEXDC; 1. SMART; SM00490; HELICC; 1.
                                                                                                                                                                                                                                                                                                                      STRAIN-BRISTOL N2;
Fulton R., Wohldmann P.;
                                                                                                                                                                                                                                                                          Nature 368:32-38(1994).
                                                                                                                                                                                                                                                                                                          SEQUENCE FROM N.A.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  STRAIN-BRISTOL N2;
                                                                                                                                                                                                                                                                                                                                                                                                                     STRAIN=BRISTOL N2;
                                                                                                                                                                                                                                                                                                                                                                                                       SEQUENCE FROM N.A.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   SEQUENCE FROM N.A.
                                 SEQUENCE FROM N.A.
 NCBI_TaxID=6239;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               235
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         q
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            qq
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 Qγ
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              Db
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              δλ
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            ŏ
   οy
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    ò
```

```
602 PKKKKPRNRNERLEKAKEDFRKYLQNPGPDMVVCDEAHKLKNDDSALS-KCMVKILTKRR 660
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              509 SGMKRRAHSDRVED-FPPCDPTMD--LMIEAKEKEQAVFELCRRYELQNPPCPL---- 559
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   84 KKENETELANISGPHKKSTS -- TSTRKRARSSKKKATDSVSDKIDESVASYDSSTHLRR 140
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      87 KSESESDESDEEEDRKKSKSKKKVDQKKKEKSKKKRTTSSSEDEDSDEEREQKSK---KK 143
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         141 SSRSKKPVNYNSSSESEEGI--SKATKKVKQKEEEEYVEEVDE------KS 185
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     -----EFEPVVPEQLETPISKRRRSRSSAKNLEKE 223
                                                                      EILSRMKLNEQLNKDAVLITHLGGTFEGKKETL--------DRFR----- 417
                                                                                                                      551 EELDTIDVNE------LDSYKTIEDRRRALKAWHSSKTPSVMIIGYDLFRILTVEDD 601
                                                                                                                                                                                                           418 -----KNYQRLSDSVKARL--------VLENDDVSWSVQDLLPLCQELN 453
                                                                                                                                                                                                                                                                                                                                                     454 I-----PLVLDWHHHNIVPGTLREGSLDLMPLIPTIRETWTRKGITQKQHYSESADPTAI 508
                                                                                                                                                                                                                                                                                                                                                                                                                    661 ICLTGTPLQNNLMEYHCMVNFVKPG-----LLGTKTEFANRFVNIINRGRTKDASPLEV 714
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             715 SFMKRRCHV -- LYDHLKKCVDRKDYRVLTEAIPPKQEYVINVROTERQ --- CALYNAFLN 769
491 GGILAHCMGLGKTLQVITFLHTVLMHEKIGEKCKRVLVVVPKNVIINWFKEFQKWLVDND 550
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 48; Gaps
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        Eukaryota; Metazoa; Nematoda; Chromadorea; Rhabditida; Rhabditoidea;
Rhabditidae; Paloderinae; Caenorhabditis.
NCBL_TaxID=6239;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           Villard L., Fontes M., Ewbank J.J.;
"Characterization of xnp-1, a Caenorhabditis elegans gene similar to
the human XNPATR. y gene.";
Gene 236:13-19(1999).
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  770 DIVGDSGLSKRLLPDYHMFSRIWTHPYQLVLHEQRM--ERERVMREDAEEE 818
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      Query Match

4.8%; Score 148; DB 5; Length 1359;
Best Local Similarity 27.1%; Pred. No. 0.38;
Matches 52; Conservative 29; Mismatches 63; Indels 41
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   560 EIMG-----PEYDQ-TRDGYYPPGA---EKRLTARKRRSRKEEVEED 597
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          EB4342547D4F4E64 CRC64;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        01-MAY-2000 (TrEMBLrel. 13, Last sequence update) 01-JUN-2001 (TrEMBLrel. 17, Last annotation update)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 -i- SIMILARITY: TO HELICASE C-TERMINAL DOMAIN.
EMBL; AF134186; AAD55361.1; -.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            1359 AA
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           01-MAY-2000 (TrEMBLrel. 13, Created)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            SEQUENCE FROM N.A. MEDLINE=99365296; PubMed=10433961;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      156191 MW;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       Interpro; IPR001410; DEAD.
Interpro; IPR001650; Helicase_C.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      InterPro; IPR000330; SNF2_N.
Pfam; PF00271; helicase_C: 1.
Priam; PF00176; SNF2_N: 1.
SMART; SM00487; DEXDC; 1.
SMART; SM00490; HELICC; 1.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            PRELIMINARY;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     Caenorhabditis elegans.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            ATP-binding; Helicase.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          1359 AA;
```

7;

Db 259 SESEKSDEEEEE 270

Search completed: January 15, 2002, 14:06:13 Job time: 952 sec

		·	

GenCore version 4.5 Copyright (c) 1993 - 2000 Compugen Ltd.

OM protein - protein search, using sw model

January 15, 2002, 13:52:18; Search time 74.94 Seconds Run on:

(without alignments)
608.868 Million cell updates/sec US-09-724-296-2_COPY_230_828 Title:

....KRLTARKRRSRKEEVEEDEK 599

BLOSUM62 Gapop 10.0 , Gapext 0.5 Scoring table:

1 MLRLLKRNIQISKRIVFTIL.....

Perfect score:

Sedneuce:

219241 seqs, 76174552 residues Searched:

Total number of hits satisfying chosen parameters:

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Listing first 45 summaries Maximum Match 100%

pir1:* pir2:* pir3:* PIR_68:* Database

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

UV-endonuclease hypothetical prote probable UV damage probable thioredox hypothetical prote translation intiat hypothetical prote polymerase-associa hypothetical prote protein kinase (EC erythrocyte membra hypothetical prote rhoptry-associated polymerase-associa oviduct-specific p major merozoite su protein F1504.11 [hypothetical prote myosin-likė protei natural killer cel major merozoite su UV-endonuclease nucleolar protein skeletal myosin nucleolin homolog hypothetical Description SUMMARIES C75350 S55785 T40373 T34036 A59294 S38173 SAZQK1 S72668 S46769 T00365 S04090 S27833 RRNZ83 T38786 S48326 A47328 T14050 T09648 T15264 E83974 RRNZ73 T41987 D71623 S05603 % Query Match Length DB 1957 322 326 500 726 911 1359 782 568 918 1403 568 1205 1702 480 635 2441 1631 100.0 Score 360 154 148.5 148.5 147.5 147. 140.5 139.5 139.5 138.5 138.5 137.5 136.5 229.5 142.5 1011. 306 Result

prote

	30	134.5	4.3	1640	2	A24594		orobable major enr
	31	134	4.3	2105	,-	A44059		genome nellinger sur
	32	133.5	4.3	1336	~	541794		genome polyprocein
	33	133	4.3	1098	~	T08599		probable transcrip
	34	132.5	4.3	1332	7	S41552		probable transcrip
	35	132	4.2	414	~	A39205		product classing
	36	132	4.2	772	2	I50463		protein kinase - c
	37	1.32	4.2	1166	7	T29009		÷
	38	132	4.2	1665	~	T29008		hypotherical prote
	39	132	4.2	1790	7	867593		transport protein
	40	132	4.2	1827	7	T16270		hypothetical prote
	41	132	4.2	5327	7	T13564		microtubule-associ
	42	131.5	4.2	952	2	T18837		hypothetical prote
	43	131	4.2	761	7	866719		
	44	131	4.2	1353	7	T00347		
	45	130.5	4.2	853	7	T51505		
								٦
						ALIG	ALIGNMENTS	A
	RESULT	н						
	871134							
	UV-endoi	nuclease	- fissi	on year	st	(Schizos	UV-endonuclease - fission yeast (Schizosaccharomyces pombe)	
	C; Speci.	C. Species: Schizosaccharomyces pombe	osaccha	romyce	ğ,	ombe		
<	C;Date:	21-Apr-1	.997 #SE	equence.	re	vision 09	C;Date: 21-Apr-1997 #Sequence_revision 09-May-1997 #text_change	lange 20-Jun-2000
<u>/</u>	C; Acces:	C; Accession: S/II34; T398I5	134; T3	9815 5		:		
	N. takao	A, takau, M.; Tottemasu, K.; Tamamoto, K.; Nicleic Acids Bos 24 1262-1271 1666	ellasu,	1267-1	mam 171	350, K.;	Yasuı, A.	
<u>ا</u>	ATTETE	Activities Characterization of	051931	120/-T	7/5	12/1, 1996	9	
	A; Refere	A;Reference number: S71134; MUD:96188860	er: S71	134; M	JID	. 9618886(rease gene trom the	e Ilsson yeast Schizosacch
`	A; Acces:	A; Accession: S71134	1.34					
	A;Status	A; Status: preliminary; nucleic acid	inary;	nuclei	ă		sequence not shown	
 >	A; Molect	A; Molecule type: mRNA	mRNA					
	A; Residues:	ues: 1-59	1-599 <tak></tak>					
	A;Cross	referenc	es: EMB	L:D785	71;	NID:9139	A;Cross-references: EMBL:D78571; NID:g1399000; PIDN:BAA11415.1;	5.1; PID:g1399001
	A: Expers	A: Experimental source: strain sp972	Ource.	strain	Š	27.2		

cch

ó; 61 TL.SMLPQVNIGANSFSAETPVDLKKENETELANISGPHKKSTSTSTRKRARSSKKKATDS 120 61 TLSMLPQVNIGANSFSAETPVDLKKENETELANISGPHKKSTSTSTRKRARSSKKRATDS 120 DKPIPWRGRLGYACLNTILRSMKERVFCSRTCRITTIQRDGLESVKQLGTONVLDLIKLU 300 VSDKIDESVASYDSSTHLRRSSRSKKPVNYNSSSESEEGQISKATKKVKQKEEEEYVEE 180 VDEKSLKNESSSDEFEPVVPEQLETPISKRRSRSSAKNLEKESTMNLDDHAPREMFDCL 240 0; Gaps 1 MLRILKRNIQISKRIVFTILKQKAFKGNHPCVPSVCTITYSRFHCLPDTLKSLLPMSSKT 60 DB 2; Length 599; Indels . 100.0%; Score 3112; DB 2; 100.0%; Pred. No. 2e-163; tive 0; Mismatches 0 Matches 599; Conservative Best Local Similarity A; Map position: Query Match 181 121 241 οy qq qq δŏ $Q_{\underline{Y}}$ g ŏ

myosin heavy chain

hypothetical

A:Molecule type: DNA A:Residues: 1-599 <LYN> A:Cross-references: EMBL:AL023859; PIDN:CAA19577.1; GSPDB:GN00067; SPDB:SPBC19C7.09c A:Experimental source: strain 972h-; cosmid c19C7

C;Genetics: A;Gene: UVDE; SPBC19C7.09c

A;Experimental source: strain SP972
R;Lyne, M.; Rajandream, M.A.; Barrell, B.G.; Lucas, M.; Gaillardin, C. submitted to the EMBL Data Library, June 1998
A;Reference number: 221881
A;Accession: T39815

A; Status: preliminary; translated from GB/EMBL/DDBJ

Qy 427 VKARLVLENDDVSWSVQDLLPLCQELNIPLVLDWHHHNIVP	RESULT 3 S55418 UV-endonucLease homolog ywjD - Bacillus subtilis C;Species: Bacillus subtilis C;Species: Bacillus subtilis C;Accession: 555418; B70060 R;Glaser, P.; Danchin, A. Submitted to the EMBL Data Library, May 1995 A;Description: Cloning and sequencing of the Bacillus A;Reference number: 555414 A;Accession: 555418 A;A	Riger. Natural Novelta, C.; Rocha, E.; Roche, B.; Rose A. Authors: Schleich, S.; Schroeter, R.; Scoffone, F. akeuchi, M.; Tamakoshi, A.; Tanaka, T.; Terpstra. T.; Winters, D.; Wipat, A.; Tanaka, T.; Terpstra. T.; Winters, D.; Wipat, A.; Yamamoto, H.; Yamane, K. A. Authors: Yoshikawa, H.F.; Zumstein, E.; Yoshikawa, A.Tille: The complete genome sequence of the Gram-po A. Recession: B70060 A.; Reference number: A6980; MUID:98044033 A.Accession: B70060 A.; Status: nucleic acid sequence not shown; translati A; Molecule type: DNA A. Residues: 1-320 < KUN> A.; Cross-references: GB:299123; GB:AL009126; NID:9263 A.; Cross-references: GB:299123; GB:AL009126; NID:9263 A.; Genetics: A.; Genetics:	Query Match Best Local Similarity 29.8%; Pred. No. 6.6e-13; Batches 84; Conservative 61; Mismatches 109 Qy 276 TIORDGLESVKOLGTONVLDLIKLVEWNHNFGIHFMRVSSDL Qy 334 OSHILEVGKLANKYNHRLTMHPGOYTQIASPREVVVDSAIRD Qy 334 OSHILEVGKLANKYNHRLTMHPGOYTQIASPREVVVDSAIRD Qy 334 OSHILEVGKLANKYNHRLTMHPGOYTQIASPREVVVDSAIRD Qy 334 OSHILEVGKTRANKYNHRLTMHPGOYTQIASPREVVVDSAIRD Qy 334 OSHILEGTFECKKTLANKYNHRLTMHPGOYTQIASPREVVVDSAIRD Qy 334 OSHILEGTFECKKTLANKYNHRLTMHPGOYTQIASPREVVVDSAIRD Qy 334 OSHILEGTFECKKTLANKYNHRLTHTHTHTHTHTHTHTHTHTHTHTHTHTHTHTHTHTHT
Db 241 DKPIPWRGRLGYACLNTILRSMKERVECSRTCRITTIQRDGLESVKQLGTQNVLDLIKLV 300 9y 301 EWNHNFGIHFWRVSSDLFPFASHAKYGYTLEFAQSHLEEVGKLANKYNHRLTWHPGQYTQ 360	541 QAVEELCRRYELQNPPCPLEIMGPEYDOTRDGYYPPGAEKRLTARKRRSRKEEVEEDEK 599 541 QAVEELCRRYELQNPPCPLEIMGPEYDOTRDGYYPPGAEKRLTARKRRSRKEEVEEDEK 599 SULT 2 5262 -endonuclease - Neurospora crassa Accession: 555262 Pating H.; Takao, M.; Yasuhira, S.; Zhao, J.H.; Ishii, C.; Inoue, H.; Yasui, A. BO J. 14, 2393-2399, 1995 Title: A eukaryotic gene encoding an endonuclease that specifically repairs DNA damagnerence number: 55562; MUID:9529280 Accession: 555262; MUID:9529280 Accession: 55562; MUID:9529280 Reference number: 55562; MUID:9529280 Status: nucleic acid sequence not shown Molecule type: mRNA Residues: 1-656 - 47AJ> Cross-references: EMBL:D11392; NID:91526560; PID:9927215 Genetics:	Ouery Match Best Local Similarity 36.6%; pred. No. 3.2e-48; Matches 241; Conservative 88; Mismatches 149; Indels 181; Gaps 18; Qy 110 ARSSKKATDSVSDKIDESVASYDSSTHLRRSSRSKRPVNVNSSESE- 157 :	Qy 233PREMEDCLDKPIDWRGRLGYACLNTILRSMKERVFCSRTCRIT 275 Db 174 NIIEPEDAQDAAERGAARPPAVNSSYLPLPWKGRLGYACLNTYLRNSKPPIFSSRTCRMA 233 Qy 276 TI

```
ujita, Y.; Fuma, S.; Galizzi, A.; Gal; Holsappel, S.; Hosono, S.; Hullo, M.; Kurlta, K.; Lapidus, A.; Lardino Levine, A.; Liu, H.; Masuda, S.; Mau S.H.; Parro, V.; Pohl, T.M.; Portete Rose, M.; Sadale, Y.; Sato, T.; Scanl F.; Sekiguchi, J.; Sekowska, A.; Se P.; Tognoni, A.; Tosato, V.; Uchiya P.; Tognoni, A.; Tosato, V.; Voshida awa, H.; Danchin, A.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    us subtilis chromosomal region fro
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     .CAA89865.1; PID:9853757
A.M.; Alloni, G.; Azevedo, V.; Ber
II., B.; Capuano, V.; Carter, N.M.;
Irington, J.; Fabret, C.; Ferrari,
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                636240; PIDN:CAB15748.1; PID:e11862
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       æ
                                                                                                                                                                                                                                                                                                                                                                                                         #text_change 15-Oct-1999
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               ENDDVSWSVQDLLPLCQELN 453
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     RDLAYHDEILSRMKLNEQLN 393
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                |:||| :| | : ::
TDMAYHYRMLEAMGIADR-- 147
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            DLFPFASH--AKYGYTLEFA 333
                                                                         3;
09; Indels 28; Gaps
                                                                                                                                                      -PGTLREGSLDLM--PLIP 482
                          GGKRKRTIDEEAAEPEEVD 592
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                SIVPLATHPDVMWDFVTPF- 89
                                                                                                                                                                                                                                  ARKRRSRK----EEVEED 597
                                                                                                                                                                                                                                                        :|| :| | | | || || || || ::
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   : ::|: | ::: | | | |
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           cion not shown
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    Length 320;
```

```
A:Residues: 1-326 <WHI>
A:Cross-references: GB:AE002022; GB:AE000513; NID:g6459590; PIDN:AAF11370.1; PID:g645
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               N;Alternate names: probable RNA-binding protein RBD18
C;Species: Schizosaccharomyces pombe
C;Date: 28-Oct-1996 *Requence_revision 07-Feb-1997 *Lext_change 17-Nov-2000
C;Date: 28-Oct-1996 *Requence_revision 07-Feb-1997 *Lext_change 17-Nov-2000
C;Accession: S55785; S68087; T37634; S52318
R;Gulli, M.P.; Girard, J.P.; Zabetakis, D.; Lapeyre, B.; Melese, T.; Caizergues-Ferre Nucleic Acids Res. 23, 1912-1918, 1995
A;Title: gar2 is a nucleolar protein from Schizosaccharomyces pombe required for 18S
A;Reference number: S55785; MUID:95319932
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  A; Residues: 1-500 <GUL>
A; Cross-references: EMBL: Z48166; NID: 9663261; PIDN: CAA88179.1; PID: 9663262
A; Cross-references: EMBL: Z48166; NID: 9663261; PIDN: CAA88179.1; PID: 9663262
Curr, Genet. 29, 307-315, 1996
A; Title: Molecular analysis of a novel Schizosaccharomyces pombe gene containing two A; Reference number: $68083; MUID: 96171513
A; Reference number: $68087
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           A;Status: nucleic acid sequence not shown; not compared with conceptual translation A;Molecule type: mRNA
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      A;Residues: 1-338,'S',340-500 <BRO>
A;Cross-references: EMBL:299091; PIDN:CAB11772.1; GSPDB:GN00066; SPDB:SPAC13F5.09
A;Experimental source: strain 972h-; cosmid c13F5
C;Genetics:
                                                                                                                                                                                                                                                                                                                                                                         10;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     89 RLYRLSSSLFPMLDLAGDDTGAAVLTHLAPQLLEAGHAFTDAGVRLLMHPEQFIVLNSDR 148
                                                                                                                                                                                                                                                                                                                                                                                                                      249 RLGYACLNTILRSMKERVFCSRTCRITTIQRDGLESVKQLGTQNVLDLIKLVEWNHNFGI 308
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     309 HFMRVSSDLFP---FASHAKYGYTLEFAQSHLEEVGKLANKYNHRLTMHPGQYTQIASPR 365
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     EVVVDSAIRDLAYHDEI----LSRMKLNEQLNKDAVLIIHLGGTFEGKKETLDRFRKNY 420
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            | :| :: | : | : | : | 149 PEVRESSVRAMSAHARVMDGLGLARTPWN-----LLLLLHGGKGGRGAE-----LAALI 196
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       QRLSDSVKARLVLENDDVSWSVQDLLPLCQELNIPLVLDWHHHNIVPGTLREGSLDLMP- 479
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             197 PDLPDPVRLRLGLENDERAYSPAELLPICEATGTPLVFDAHHHVV------HDKLPD 247
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   480 -LIPTIRETWTRKGIT-----QKQHYSESADPTAISGMKRRAHSDRVFDFPPCDPTMD- 531
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             248 QEDPSVREWVLRARATWQPPEWQVVHLS----NGIEGPQDRRHSHLIADFPSAYADVPW 302
                                                                                                                                                                                                                                                                                                                                                                         Gaps
                                                                                                                                                                                                                                                                                                                                                                                                                                                       A;Residues: 374-407 <VAN>
NS. Brown, D.: Churcher, C.M.: Barrell, B.G.; Rajandream, M.A.: Wood, Submitted to the EMBL Data Library, September 1997
A;Reference number: Z21733
                                                                                                                                                                                                                                                                                                                Length 326;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      nucleolar protein gar2 - fission yeast (Schizosaccharomyces pombe)
                                                                                                                                                                                                                                                                                                                                                                      Indels
                                                                                                                                                                                                                                                                                                       7.4%; Score 229.5; DB 2; 27.9%; Pred. No. 9.6e-06; tive 42; Mismatches 140;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            A; Status: preliminary; translated from GB/EMBL/DDBJ
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           A;Status: nucleic acid sequence not shown A;Molecule type: DNA
                                                                                                                                 A; Experimental source: strain R1
                                                                                                                                                                                                                                                                                                                                                             88; Conservative
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      532 LMIEAKEKEQAVFEL 546
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            303 IEVEAKGKEEAIAAL 317
                                                                                                                                                                                                                                                                                                                                     Best Local Similarity
               A;Status: preliminary
                                                A; Molecule type: DNA
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       A; Molecule type: DNA
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 A; Accession: T37634
                                                                                                                                                                                                                  A; Map position: 1
                                                                                                                                                      C;Genetics:
A;Gene: DR1819
                                                                                                                                                                                                                                                                                                       Query Match
                                                                                                                                                                                                                                                                                                                                                             Matches
                                                                                                                                                                                                                                                                                                                                                                                                                         δ
                                                                                                                                                                                                                                                                                                                                                                                                                                                                         qq
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     δ
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           QΩ
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     δy
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        ОD
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   Qγ
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       qq
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         QQ
                                                                                                                                                                                                                                                                                                                                                                                                                                                          C;Accession: E83974
R;Takami, H.; Nakasone, K.; Takaki, Y.; Maeno, G.; Sasaki, R.; Masui, N.; Fuji, F.; Hira Nucleic Acids Ros. 28, 4317-4331, 2000
A;Title: Complete genome sequence of the alkaliphilic bacterium Bacillus halodurans and A;Reference number: A83650; MUID:20263314
A;Accession: E83974
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              R.J.;
C.; Ma
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          A;Cross-references: GB:AP001516; GB:BA000004; NID:g10175192; PIDN:BAB06316.1; GSPDB:GN0C
A;Experimental source: strain C-125
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 A;Title: Genome sequence of the radioresistant bacterium Deinococcus radiodurans R1. A;Reference number: A75250; MUID:20036896 A;Accession: C75350
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         probable UV damage endonuclease - Deinococcus radiodurans (strain R1) ("Species: Deinococcus radiodurans C'Species: Deinococcus radiodurans ("Date: 03-Dec-1999 #sequence_revision 03-Dec-1999 #text_change 31-Mar-2000 C;Accession: C75350 R;White, O.; Elsen, J.A.; Heldelberg, J.F.; Hickey, E.K.; Peterson, J.D.; Dodson, M.; Shen, M.; Vamathevan, J.J.; Lam, P.; McDonald, L.; Utterback, T.; Zalewski, Science 286, 1571-1577, 1999.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     6
                                                                                                                                                                                                                                                                                                                                                                                                       C.Species: Bacillus halodurans
C.Date: 01-Dec.2000 #sequence_revision 01-Dec-2000 #text_change 31-Dec-2000
                                                                                                                                                                                                                                                                                                                                                                               protein BH2597 [imported] - Bacillus halodurans (strain C-125)
                                                                                                            206 VPFVFDFHHFYANP----DDHADLNVALPRMIKTWERIGLOPKVHLSSPKSEQAI---- 256
--SVINIHIGGAYGNKDTATAQFHQNIKQLPQEIKERMTLENDDKTYTTEETLQVCEQED 205
                                                         454 IPLVLDWHHHNIVPGTLREGSLDLMPLIPTIRETWTRKGITQKQHYSESADPTAISGMKR 513
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               247 RGRLGYACLNTILRSMKERVFCSRTCRITTIQR-----DGLESVKQLGTQNVLDLIKLVE 301
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              302 WNHNFGIHFMRVSSDLFPFASH---AKYGYTLEFAQSHLEEVGKLANKYNHRLTMHPGQY 358
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     359 TQIASPREVVVDSAIRDLAYHDEILSRMKLNEQLNKDAVLIIHLGGTFEGKKETLDRFRK 418
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     419 NYQRLSDSVKARLVLENDDVSWSVQDLLPLCQELNIPLVLDWHHHNIVPGTLREGSLDLM 478
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            479 PLIPTIRETWTRKGITQKQHYSESADPTAISGMKRRAH-----SDRVFDF-----PPCDP 528
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    37; Gaps
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           2 RIQFGYVAMSMELANASP----SKTWTATQFEKIEDHEAGLRKLERIAKTNLHNCLRLLK 57
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    Length 322;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     Indels
                                                                                                                                                                                                                  514 RAHSDRV---FDFPPCD-----PTMDLMIEAKEKEQAVFEL 546
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  9.8%; Score 306.5; DB 2; 26.6%; Pred. No. 5.7e-10; tive 75; Mismatches 123;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               529 TMDLMIEAKEKEQAVFELCR 548
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                : :|||||:|: |:|:| :
283 HLHVMIEAKKKDLALFQLMK 302
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  Conservative
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 A; Molecule type: DNA
A; Residues: 1-322 <STO>
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 Best Local Similarity
Matches 85; Conserv
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            Status: preliminary
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              A; Gene: BH2597
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     Query Match
                                                                                                                                                                                                                                                                                                                                                                               hypothetical
                                                                                                                                                                                                                                                                                                                      RESULT
g
                                                         ò
                                                                                                               a
                                                                                                                                                                        δ
                                                                                                                                                                                                                           q
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   q
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        a
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    q
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             ô
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            ð
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               g
```

A; Gene A; Map C; Supe F; 264 - F; 367 -	A;Gene: gar2; SPDB:SPAC13F5.09 A;Map position: 1 C;Superfamily: unassigned ribonucleoprotein repeat-containing proteins; ribonucleoprotei F;264-331/Domain: ribonucleoprotein repeat homology <rrm1> F;367-433/Domain: ribonucleoprotein repeat homology <rrm2></rrm2></rrm1>	0 y 0 DP	390EVI 359 TQLAS ::
Query Ma Best Loc Matches	Ouery Match 4.9%; Score 154; DB 2; Length 500; Best Local Similarity 27.9%; Pred. No. 0.23; Matches 53; Conservative 33; Mismatches 60; Indels 44; Gaps 6;	Qy	396 AVLI:: :: :: 508 SVVL
Qy Db	48 DTLKSLLPMSSKTTLSMLPQVNIGANSFSAETPVDLKKENETELANISGPHKKSTSTSTR 107 :	Qy	441 SVQD)
Qy Db	108 KRARSSKKKATDSVSDKIDESVASYDSSTHLRRSSRSKKPVNYNSSESESEBQISKATK 167 :	Qy	494 TQKQI 619 AQDI
Qy	168 KVKQKEEBEYVEEVDEKSLKNESSSDEFEPVVPEQLETPISKRRRSRSS 216 : :	RESULT S51441	88 3
QY Db	217 AKNLEKESTM 226 ::: : :: 182 SESEBEEVV 191	N; A1 C; Sp C; Da	N; Alternate nan C; Species: Sac C; Date: 23-Feb C; Accession: S
RESULT T40373 probab	7 le thioredoxin - fission yeast (Schizosaccharomyces pombe)	R;Pa subm A;De A;Re	
C;Spec C;Date C;Acce R;Wood	<pre>ies: Schizosaccharomyces pombe : 03-bec-1999 #sequence_revision 03-bec-1999 #text_change 03-bec-1999 ssion: T40373 . V : Rajandream, M.A.; Barrell, B.G.; Hilbert, H.; Duesterhoeft, A.</pre>	A; AC A; MO A; Re A; Cr	A,Accession: 5 A,Molecule typ A,Residues: 1- A,Cross-refere
submit A;Refe A;Acce A;Stat	ted to the EMBL Data Library, December 1998 rence number: 221923 ssion: T40373 us: preliminary; translated from GB/EMBL/DDBJ	C;Genet: A;Gene: A;Cross: A;Map_po	C,Genetics: A,Gene: SGD:IM A,Cross-refere A,Map position
A; Mole A; Resi A; Cros A; Expe C; Gene	A; Molecule type: DNA A; Residues: 1-726 < WOO> A; Residues: 1-726 < WOO> A; Cross-references: EMBL: 295620; PIDN: CABO9115.1; GSPDB:GN00067; SPDB: SPBC3D6.13c A; Experimental source: strain 972h-; cosmid c3D6 C; Genetics:	Qu Ma	Query Match Best Local S Matches 113
A;Gene A;Map	: SPDB:SPBC3D6.13c position: 2	QY	6 KRNI :
Quer	Query Match	qa o	202 KEEV
Best Lo Matches	Local Similarity 22.3%; hes 127; Conservative 70	à a	
Qy Dp	44 HCLPDTLKSLLPMSSKTTLSMLPQVNIGANSFSAETPVDLKKENETELANI 94	δλ	
Qy Db	95 SGPHKKSTSTSTRKRARSSKKKATDSVSDKIDESVASYDSSTHLRRSSRSKKPVNVNSSS 154 : : :	gg Åð	
yo g		Dp Qy	337 KDCE 214
Cy Oy	SSAKNLEKESTWNLDDHAPREMFDCLDKPIPWRGRLGYACLNT	Db Qy	
δy d		Db Qy	445 YEAK 305 NFGI
o v	RVSSDLFPFASHAKYGYTLEFRAGSHLEEVGKLANKYNHRLTMHPGQY 35	qq	493 QLTY

```
nces: EMBL:U17247; NID:9577216; PIDN:AAB67359.1; PID:9577221; MIPS:YLR3
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  21;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         rotein YLR309c - yeast (Saccharomyces cerevisiae)
mes: hypothetical protein L2142.5
charomyces cerevisiae
-1995 #sequence_revision 12-May-1995 #text_change 29-Oct-1999
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               KVKQKEEEEEYVEEVDEKSLKN---ESSSDEFEPVVPEQLETPISKRRRS----- 213
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           EKSNITKELTSLRTSYKQKEKTVSYLBEQVKQFSEQKDVAEKSTEQLRKDHAKISN 552
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      KNKGGITTGDISEEETVDNSINTEE------YDKLKENLOELQEKY 336
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 HEMRVSSDLFPFASHAK------YGYTLEFAQSHLEEVGKLANKYNH 349
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         4.8%: Score 148.5; DB 2; Length 911;
illarity 20.4%; Pred. No. 1;
Conservative 75; Mismatches 180; Indels 185; Gaps
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               LPQVNIGANSFSAETPVDLKKENETELANISG------PHKKSTSTSTRKRAR 111
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  LAELNI------TLKEKERKLSELEKKMKELPKAISHONVGNNNRRKKNR 292
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        K-----ATDSVSDK--IDESVASYDSSTHLRRSSRSKKPVNYNSSSESESE-EQIS 163
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           ------RSSAKNLEKEST--MNLDD--HAPREMFDCLDKPIPWRGRLG 251
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                TVGNELVDAKDEIKESSSKQNEEVKTVKLELDDLRHKNATMIE------A 444
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               NTILRSMKERVFCSRTCRITTIQRDGLESVKQLGTQNVLD-----LIKLVEWNH 304
                                                                                                                                                                                                                     SPREVVVDSAI-----EQLNKD 395
                                                                                                                             FPKLIVVRDGIASYYPAKMAQDNKDYRRILGWMKNNWLPVLPELRTSNSKEIFNDE 507
                                                                                                                                                                                   IHLGGTFEGKKET------LDR----FRKNYQRLSDSVKARLVLENDDVSW 440
                                                                                                                                                                                                                                                                                          LLPLCQELNI-----PLVLDWHHHNIVPGTLREGSLDLMPLIPTIRETWTRKGI 493
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         SISE-----LKEELAIKNHSLEDSRMKITELEQN------LSSKSTIMEE 244
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      QISKRIVFTILKQKAFKGNHPCVPSVCTITYSRFHCLPDTLKSLLPMSSKTTL--- 62
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   The sequence of S. cerevisiae cosmid L2142. mber: S51437
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               he EMBL Data Library, November 1994
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   nces: SGD:S0004300; MIPS:YLR309c: 12R
                                                                                                                                                                                                                                                                                                                                                                                                            HYSESADPTAISGMKRRAHSDRVFD 522
                                                                                                                                                                                                                                                                                                                                                                                                                                                 YWDETAKGTPIS----IEKDTVFD 642
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              imilarity
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     911 <PAU>
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    51441
```

oy Bb	350 RLTMHPGQYTQIASPREVVVDSAIRDLAYHDEILSRWKLNEQLNKDAVLIIH 401	
	LGGTFEGKKETLDRFRKNYQRLSDSVKARLVLENDDSWSVQDLLFLCQELNIPLVLDWH 46	
Qy Dp	462 HHNIVPGTLREGS 474 :	
RESULT	o :	
T34036 hypoth C;Spec	s elega	
C; Date C; Acce	<pre>e: 29-0ct-1999 #sequence_revision 29-0ct-1999 #text_change 29-0ct-1999 ession: T34036 #CAT D : #CATIONS P</pre>	
Submit Submit A; Desc A; Refe	Submitted to the EMBL Data Library, April 1997 A; Description: The sequence of C. elegans cosmid B0041.	
A;Acce A;Stat A;Mole	##.Accession: Tayuso A.Status: preliminary; translated from GB/EMBL/DDBJ A.Wolecule type: DNA	
A; Resi A; Cros	1-135 erence	
C; Genet A; Genet	crone	
A;Mdp A;Intr	A,Tatp position: 1 A,Introns: 12/2; 59/2; 248/2; 582/1; 675/1; 733/3; 966/2; 1044/2; 1310/3	
Quer Best Matc	Query Match Best Local Similarity 20.9%; Pred. No. 1.7; Matches 136; Conservative 96; Mismatches 222; Indels 197; Gaps 34;	
oy Ob	83 LKKENETELANISGPHKKSTSTSTRKRARSSKKKATDSVSDKIDESVASYDSSTHLRRSS 142 : :	
	RSKKPVNYNS-SSESSESEQISKATKKVKQKEEEEYVEEVDEKS:	
Oy Db	18f LKNESSSDEFEPVVPEQLETPISKRRRSRSSAKNLEKESTMNLDDHAPR 234 ::	
oy Dp	235 EMFDCLDKPIPWRGRLGYACLNTILRSMKERVFCSRTCRITTIGRDGLE 283	
qq	449 SKKPVEVHNSLVRILKPHQAHGIQFMYDCAFESLDRLDTEGS 490	
Oy Dp	335SHLEEVGKLANKYNHRLTWHPGQYTQIASPREVVVDSAIRDLAYHD 380 : : : : : : : : : : : : :	
Οy	381 EILSRMKLNEQLNKDAVLIIHLGGTFEGKKETLDRFR 417	
QQ	551 EELDTIDVNELDSYKTIEDRRRALKAWHSSKTPSVMIIGYDLFRILTVEDD 601	
	KNYQRLSDSVKARL	
	PKKKKPRNRNRRLEKAKEDFRKYLQNPGPDMVVCDEAHKLKNDDSALS-KCMVKILTKRR	
λά	454 IPLVLDWHHHNIVPGTLREGSLDLMPLIPTIRETWTRKGITQKQHYSESADPTAI 508	

```
Riddey, R.G.; Takacs, B.; Lahm, H.W.; Delves, C.J.; Goman, M.; Certa, U.; Matile, H. Mol. Biochem. Parasitol. 41, 125-134, 1990
A;Title: Characterisation and sequence of a protective rhoptry antigen from Plasmodiu A;Teference number: A45514; MUID:90348711
A;Steference number: A45514
A;Status: preliminary
A;Status: preliminary
A;Residues: 1-782 < RI2>
A;Cross-references: GB:M32853; NID:g160656; PID:g160657
C;Superfamily: Plasmodium falciparum rhoptry-associated protein 1
                                                                                                                                                                                                                                                                                              - malaria parasite (Plasmodium falciparum)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     39;
                                                                                                                                                                                                                                                                                         rhoptry-associated protein 1 precursor - malaria parasite (Plasmodium falcip
N:Alternate names: protective antigen
C;Species: Plasmodium falciparum
C;Date: 17-Apr-1993 #sequence_revision 17-Apr-1993 #text_change 09-Jun-2000
C;Accession: A45514; S27833
                                                                      SGMKRRAHSDRVFD-FPPCDPTMD--LMIEAKEKEQAVFELCRRYELQNPPCPL----- 559
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             43 FHCLPDTLKSLLPMSSKTTLSMLPQVNIGANSFSAETPVDLKKENETELANISGPHKKST 102
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            FKASPSVVKTSTPSGTQT-----SGLKSSSPSSTKSSSPSNVKSAS-PHGESN 168
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     STSTRKRARSSKKKATDSVSDKIDESV -- ASYDSSTHLRRSSRSKKPVNYN -- -- SSSES 156
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         365 REV------KDAVLII- 400
661 ICLTGTPLQNNLMEYHCMVNFVKPG-----LLGTKTEFANRFVNIINRGRTKDASPLEV 714
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               S-SEESTTKSSKRSASVAGIVGADEEAPPAPKNTLTPL----EELYPINVNLFNYKYSLN 223
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       224 NMEENINILKNEGDLVAQKEEFEYDENMEKAKQDKKKALEKIGKQSDEEPFMFSENKFLE 283
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                284 NOVKERNVAGSFSRFFSKLNPFKKDEVIEKTEVSKKTFSGIGFNLTDKEAKVLGVGATYQ 343
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                ----RLGYACLNT-----ILRSMKERV--FCSRTCRITTIQRDGLESVKQLGTQNVLDLI 297
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       298 KLVEWNH-----GYTLE----FAQS 335
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               336 H------LEEVGKLAN------KYNHR------LTMHPGQ--YTQIASP 364
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   401 -- HLGGTFEGKKETLDRFRKN-----YQRLSDSVK-----ARLVLENDDVSW---SV 442
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          QDELPLCQELNIPLYLDWHHHNIVPGTLREGSLDLMPLIPTIRETWTRKGITQKQHYSES 502
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       ----EKEIKPAKKYSTS 669
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              ESEEQISKATKK---VKQKEEEEFYVEEVD-----EKSLKNESSSDEFEPVVPEQ---LE
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             205 TPISKRRSRSSAKNLEKESTMNLDDHAPREM------FDCLDKPIPWRG-----
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        DB 2; Length 782;
                                                                                                                                                                :1:1 :1 :1 :1 :1 | 1:1: | 1:1: | 1:1: | DIVGDSGLSKRLLPDYHMFSRIWTHPYQLVLHEQRM--ERERVMREDAEEE 818
                                                                                                                                    EIMG----- -- PEYDQ-TRDGYYPPGA---EKRLTARKRRSRKEEVEED 597
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  Mismatches 241;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   4.7%; Score 147.5; DE 20.8%; Pred. No. 0.94;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         635 KHFLKLVPRHN--FLLDYHFNSIF-----
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              95;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           Best_Local Similarity 20.8 Matches 146; Conservative
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        Query Match
                                            509
                                                                                                                                                                              770
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        122
                                                                                       715
                                                                                                                                  560
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       103
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 169
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            157
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             249
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  400
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            457
                                                                                     Dp
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        QQ
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    Op
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      qq
                                          δ
                                                                                                                                    ò
                                                                                                                                                                          g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               ó
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       á
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               qq
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            δò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    \delta
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             ŏ
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     δy
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               q
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          Οy
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         ŏ
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   qq
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                δλ
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             g
```

```
A; Molecule type: DNA
A; Residues: 1-113 <BOW>
A; Residues: 1-13 <BOW>
A; Cross-references: EMBL: 238114; NID: 9558402; PIDN: CAA86249.1; PID: 9558403; MIPS: YMLO
R; Brown, D.: Bowman, S.
submitted to the EMBL Data Library, October 1994
                                                                                                                                                                                                                                                                                                                                             A;Accession: T38056
A;Status: preliminary; translated from GB/EMBL/DDBJ
A;Molecule type: DNA
A;Molecule type: DNA
A;Residues: 1-401 <SX2>
A;Cross_references: EMBL:Z98599; PIDN:CAB11250.2; GSPDB:GN00066; SPDB:SPACIEII.01c
A;Cross_references: strain 972h-; cosmid clE11
C;Genetics:
                                                                                                                                                                                                                                           A;Molecule type: DNA
A;Residues: 280-913 <SKE>
A;Cross-references: EMBL:Z98762; PIDN:CAB11485.1; GSPDB:GN00066; SPDB:SPAC4A8.16c
C;Date: 15-Sep-2000 #sequence_revision 15-Sep-2000 #text_change 15-Sep-2000 C;Accession: T38786; T38056 R;Skelton, J; Churcher, C.M.; Barrell, B.G.; Rajandream, M.A.; Wood, V. submiited to the EMBL Data Library, August 1997 A;Reference number: Z21751
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    C;Species: Saccharomyces cerevisiae
C;Date: 02-Dec-1994 #sequence_revision 09-Mar-1996 #text_change 29-Oct-1999
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      71 GANSFSAETPVDLKKENETELANISGPHKKSTSTSTRKRARSSKKKATDSVSDKIDESVA 130
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            131 SYDSSTHLRRSSRSKKPVNYNSSSESEEGISKATKKVKQKEEEEYVEEVDEKSLKNES 190
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          62 SEESEVEVPK----KKAVAASEDSESDSE-----SSEEEEETESEEDSEVSDESESSES 112
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               191 SSDEFEPVVPEQLETPISKRRRSRSSAKNLEKE-----STMNLDDHAPREM 236
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                113 ESESEEESEESEE-ESDESERSGPSSFLKKPEKEEAKPAGLKFLRGESSEESSDEEEGRRV 171
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  237 F-DCLDKPIPWRGRLGYACLNTILRSMKERVFCSRTCRITTIQRDGLESVKQLGTQNVLD 295
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    296 LIKLVEWNHNFG--THFMRVSSDLFPFASHAKYGYTLEFAQSHLEEVGKLANKYNHRLT- 352
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      215 --KCKEAGRNPPPYTEFLSALDQKLESADKA-----FIKSLDAANGRAFNALKQRVRK 265
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           353 -----EVVVDS-AIRDLAYHDEILSRM 386
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            266 NNRQFQSDIDRYRKDPEGFMKPAELNEIPKPAGKAGQDEVIVDGVATRGIVAPTEGLGK- 324
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 387 KLNEQENKDAVL/ITHLGGTFE--GKKETLDRFRK-------NYQRLSDSVKARL 431
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  325 --PEEIT-PADIFKYLRAIFEARGKKST-DRSEQIRLLEKLSTIAVTDYQRL--RVKVAL 378
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  172 VKSAKDK----RYEEFISCMETIKNAMS-----SNNWIVVSNEFDHLNKVSQ-----
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 Indels 107;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                DB 2; Length 918;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            hypothetical protein YML071c - yeast (Saccharomyces cerevisiae)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 57; Mismatches 172;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            4.7%; Score 145.5; D
23.5%; Pred. No. 1.5;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          C; Accession: $48326; $48825
R; Bowman, S; Churcher, C.
Bounitted to the EMBL Data Library, August 1994
A; Reference number: $48326
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    432 VLENDDVSWSVQDLLPLCQ 450
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     379 LAVRFDINTGSGQYMPIDQ 397
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     103; Conservative
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         A;Gene: SPDB:SPACIEII.01c
A;Map position: 1
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   Query Match
Best Local Similarity
                                                                                                                                                                          A;Accession: T38786
A;Status: preliminary
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           A; Accession: S48326
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              RESULT 13
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     Matches
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            δλ
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          qq
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  δλ
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   qq
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    \delta
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      qq
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           οy
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 δŏ
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  qq
                                                                                                                                                                                                                                                                                                                                                                                                        polymerase-associated nucleocapsid phosphoprotein - parainfluenza virus type 1 (strain C ; Species: parainfluenza virus type 1 (5.Date: 30-Jun-1993 #sequence_revision 30-Jun-1993 #text_change 16-Jul-1999 (5.Accession: E40234 W.; Portner, A. Wirology 189, 340-343, 1992 A. Title: The P genes of human parainfluenza virus type 1 clinical isolates are polycistra A. Reference number: A40234; MUID:92295573 A. Title: The P genes of human parainfluenza virus type 1 clinical isolates are polycistra A. Reference number: A40234; MUID:92295573 A. Reference number: A60234 A. Residues: 1-568 <POM>
A. Residues: 1-568 <POM>
A. Residues: 1-568 <POM>
A. Residues: 1-568 <POM>
C. Stoperfamily: parainfluenza virus polymerase-associated nucleocapsid phosphoprotein C. Keywords: nucleocapsid; phosphoprotein
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             23;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       translation intiation factor eif-3 - fission yeast (Schizosaccharomyces pombe) C;Species: Schizosaccharomyces pombe
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               91 SRRTGES----KTETHARILDOQGIHRASRRGTSPNPLPENMGDERNTRIDEDSPNERRH 146
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    274 SKNEEPAVTQKNNHNHRKTKHT---TNNANNNAKCLPTPEHDTTSNEEGTSNTSVDEM-- 328
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         394 KDAVLIIHLG-----GTFEGKKETLDRFRK-----NYQRLSDSVKARLVLENDDVSWSV 442
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       329 -- AKLLVSLGVMKSQHEFELSRRASHQFAKRMLKSANYKEMTFNLCGMLLSVEKSLGNKV 386
                                    503 ---ADPTAIS----GMKRRAHSDRVFDFPPCDPTMDLMIEAKEKEQAVFELCRRYELONP 555
                                                                                                     670 HIYFDPTVASYAYYNLDRRTMVTIINDY------FEAKKKELTVIVSRMKTDM--- 716
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              62 LSMLPQVNIGANSFSAETPVDLKKENETELANISGPHKKSTSTSTRKRARSSKKKATDSV 121
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                161 QISKATKKVKQ-----KEEEETVEEVDEKSLKNESSSDEFEPVVPEQLETPISKRRSR 214
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      215 SSAKNLEKESTMNLDDHAPREMFDCLDKPIPWRGRLGYACLNTILRSMKERVFCSRTCRI 274
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              275 TTIQRDGLESVKQLGTQNVLDLIKLVEW-NHNFGIHFMRVSSDLFPFASHAKYGYTLEFA 333
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   QSHLEEVGKLANKYNHRLTMHPGQYTQIASPREVVVDSAIRDLAYHDEILSRMKLNEQLN 393
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            443 QDLLPLCQELNIPLVLDWHHHNIVPGTLREGSLDLMPLIPTIRETWTRKGITQKQHYSES 502
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             387 EENRTLLKOIQEEIDSSRDLHKRFSEYQKEQNSLAMAANLSTLHIITDRGGKT----GDP 441
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  503 ADPTAISGMKRRAHSDRV----FDFPPCDPTMDLMIEAKEKEQAVFELCRRYELQNP 555
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                442 SDTTRSPSVFTKGKDNKVKKTRF-----DPSMEALGGQEFKPDLIREDELREDIRNP 493
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            36 LSYKP-TEIGEDRSWLHNIIDNPKENKSSCKSDDNNKDRAISTSTQDH-RSSEE---SGI
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  232 TTIKRNQTRSER---TQN-SELHKSTSGDSSNLEDHNTKTSQKIPP------
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    4.7%; Score 147; DB 1; Length 568; 21.2%; Pred. No. 0.66; tive 87; Mismatches 214; Indels 122;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 122 SDKIDESVASYDSSTHLR-----RSSR---SKKPVNYNSSSE---
                                                                                                                                                                          556 PCPLEIMGPEYDQTRDGYYPPGAEKRLTARKRRSRKEEVEE 596
                                                                                                                                                                                                                                           717 ---LSLONEESKIPND----KSANSKLATRLMKKFKAEIRD 750
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 147 ORSVPTDEDRKMAENSNKREEDQVEGFPEEVRGSTSLSDDGE---
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         Matches 114; Conservative
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         Best Local Similarity
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            Query Match
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         RESULT 12
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           QΩ
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             Вb
                                    ò
                                                                                                  g
                                                                                                                                                                          óλ
                                                                                                                                                                                                                                        g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            Д
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                ŏ
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   φy
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             οy
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          qq
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               ōλ
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             qq
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            δ
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  ó
```

20;

```
A; Molecule type: DNA
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                RESULT 15
                                                                                                                                                                                                                                                                           257
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                S38173
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           qq
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     qq
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   qq
                                                                                                          q
                                                                                                                                                                                                                                                                                                                                 qq
                                                                                                                                                                                                                                                                                                                                                                                                                                         g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      a
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   q
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             δ
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 qq
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  ó
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            ò
                                                      ò
                                                                                                                                                                ò
                                                                                                                                                                                                                                                                           ò
                                                                                                                                                                                                                                                                                                                                                                                        ó
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  ŏ
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             ô
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      R;Werner, C.; Rajan, T.V.
Mol. Biochem. Parasitol. 50, 255-260, 1992
Mol. Biochem. Parasitol. 50, 255-260, 1992
A;Title: Comparison of the body wall myosin heavy chain sequences from Onchocerca volvul A;Reference number: A59294; MUID:92158005
A;Accession: A59294
                                                      'Molecule type: DNA
'Residues: 88-607 <BRO>
'Cross-references: EMBL:246373; NID:9587529; PIDN:CAA86507.1; PID:9587539; MIPS:YML071c'
'Gross-references: EMBL:246373; NID:9587529; PIDN:CAA86507.1; PID:9587539; MIPS:YML071c'
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                33;
                                                                                                                                                                                                                                                                                                     18;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    N;Alternate names: major body wall myosin
C;Species: Onchocerca volvulus
C;Date: 09-Jun-2000 #sequence_revision 09-Jun-2000 #text_change 08-Sep-2000
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           1111 LKKQVNETQNRITELEEELEN-----ERQSRSKAERAKSDLQRELEELGDRLDEQGGAT 1164
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   83 LKKE-NET-----ELANISGPHKKSTSTSTRKRARSSKKKATDSVSDKIDES---- 128
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 ----VASYDSSTHLRRSSRSKKPVNYNSSSE------SESEEQISKATKKV 169
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        316 ESSNESLIEMMVKRKIEVLREHVYMSLNVFLKSFLYDTNDLEIPFPEELESTVLRINGTN 375
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   KKSTSTSTRKRARSSKKKATDSVSDKIDESVASYDSSTHLRRSSRSKKPVNYNSSSESES 158
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             159 EEQISKAT-KKVKQKEEEETYVEEVDEKSLKNESSSDEFE--PVVPEQLETPISKRRRRSRS 215
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           216 SAKNLEKESTMNLDDHAPREMFDCLDKPIPWRGRLGYACLNTILRSMKERVFCSRTCRIT 275
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           491 -----KEEEEINKVEVTPEEP----SKKIRTSKRENKIPTNAVMLOFVDK-----C-IT 534
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     276 TIQRD---GLESVKQLGTQNVLDLI----KLVEWNHNFGIHFMRVSSDLFPFASHAKYGY 328
                                                                                                                                                                                                                                                                                                                                                                                                            258 LVKLLSTNVSVNSLKKILQYLNSIPPFDGKTNKSLLSV--FLAMRYKFITDEIASYPLDV 315
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   Gaps
                                                                                                                                                                                                                                                                                                          Gaps
                                                                                                                                                                                                                                                                                                                                                           1 MLRLLKRNIQIS--KRIVFTILKQKAFKG-NHPCVPSVCTITYSRFHCLPDTL----- 50
                                                                                                                                                                                                                                                                                                                                                                                                                                                                         51 ----KSLLPMSSKTTLSMLPQ-----VNIGANSFSAET---PVDLKKENETELANISGPH 98
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         Residues: 1.1957 < WER>
;Cross-references: GB:M74066; NID:9159892; PIDN:AAA29420.1; PID:9159893
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    Query Match
4.7%; Score 145; UB 2; Length 1357,
Best Local Similarity 18.1%; Pred. No. 4.3;
Matches 130; Conservative 111; Mismatches 231; Indels 248;
                                                                                                                                                                                                                                              4.7%; Score 145; DB 2; Length 607;
21.6%; Pred. No. 0.93;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             A; Experimental source: sex female
C; Superfamily: myosin heavy chain; myosin motor domain homology
F;87-773/Domain: myosin motor domain homology <MMO>
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        A;Status: preliminary; not compared with conceptual translation
                                                                                                                                                                                                                                                                                                        Indels
                                                                                                                                                                                                                                                                                                     73; Mismatches 160;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          skeletal myosin - nematode (Onchocerca volvulus)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  329 TLEFAQSHLEEVGKLANKY 347
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  | | :: :||:||
586 TTEQLARAIDKRAELASKY 604
                                                                                                                                                                                                                                                                                                        Conservative
A; Reference number: S48816
A; Accession: S48825
                                                                                                                                                                                                                                                                           Similarity
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     A; Molecule type: mRNA
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          Accession: A59294
                                                                                                                                                                A; Map position: 13L
                                                                                                                                                                                                                                                                                                     82;
                                                                                                                                                                                                                                                    Query Match
Best Local (
                                                                                                                                                                                                                                                                                                     Matches
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         RESULT 14
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 129
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              g
                                                                                                                                                                                                                                                                                                                                                                                                               q
                                                                                                                                                                                                                                                                                                                                                                                                                                                                         ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 a
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            a
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               δλ
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    ò
                                                                                                                                                                                                                                                                                                                                                           ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             à
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        õ
```

```
myosin-like protein MLP1 - yeast (Saccharomyces cerevisiae)
N;Alternate names: protein YKR055w; protein YKR415
C;Species: Saccharomyces cerevisiae
C;Species: Saccharomyces cerevisiae
C;Date: 03 May-1994 #sequence_revision 03-May-1994 #text_change 21-Jul-2000
C;Accession: S38173; S40647; S31207
R;Baladron, V; Ballesta, J.P.G; Bou, G; del Rey, F; Esteban, P.F.; Garcia-Cantale submitted to the Protein Sequence Database, March 1994
A;Reference number: S38158
A;Accession: S38173
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 A; Molecule type: DNA
A; Residues: 1-1875 cABAL>
A; Residues: 1-1875 cABAL>
A; Casos references: EMBL: 228320; NID: 9486586; PID: 9486587; MIPS: YKR095w
A; Cross references: EMBL: 228320; NID: 9486586; PID: 9486587; MIPS: YKR095w
A; Experimental source: strain $288C
B; Bou, G.; Esteban, P.F.; Baladron, V.; Gonzalez, G.A.; Cantalejo, J.G.; Remacha, M.; Yeast 9, 1346, 1993
A; Title: The complete sequence of a 15 820 bp segment of Saccharomyces cerevisiae chr
A; Reference number: $40644; MUID: 94205265
A; Accession: $40647
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              1623 KKLESDINELEIALD--HANQANAQAQKNVKTYQDQMRELQQQVETEQRNGRNSREQYLN 1680
                                                                                                                                                       1284 KTRLINSENTDLGKQIEEAESQVNAMTRLKAQLTSQLEEARRSLDEEARDRNNLAAQMKNY 1343
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  1387 LLKGDELEESKKRQMHKMNELQETLDAANSKISSLEKTKSRLVSDLDDAQMDVERANSYA 1446
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                1507 RENKELAQERKDLTDQLGEGGRSVFEMQKLIRRLEVEKDELQHALDQPQDALEAEESKVL 1566
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            1735 RKLDTE-----LLAIQADLDETLNEY--KASEERCKAASSDAARLAEQLRKEQENSLQND 1787
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           1447 SQLEKKQKGFDKVIDEWKKKTDDIATEVDNAQREARNVSTELFKLKSEQDEVLETIEGLR 1506
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  1681 MEKKATLLQSEKEEMSVANGQAERARKQADYD-----ANEAHTQCNELSAQAESLCGSR 1734
1165 AAQVEVNKKREAELAKLRRDLEEANMNHENQLAAIRKKHNDAVAELGDQIEQVQKAKAKI 1224
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            449
                                                                                                                                                                                                                                                                                                                                                                                                                                                                -----TILRSMKERVFCSRTCRITTIQRDGL-ESVKQLGTQNVLDLIKLVEWNHNFGIHF 310
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     A;Molecule type: DNA
A;Residues: 1-1875 <BOU>
A;Cross-references: EMBL:X73541; NID:g450550; PIDN:CAA51948.1; PID:g450554
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 398 -----LIHLGGTFEGKKETLDRFRKNYQRLSDSVKARLVLENDDVSWSVQDLLPLC
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      450 QEL----NIPLVLDWHHHNIVPGTLREG---SLDLMPLIPTIRETWTRKGITQKQHY--
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               500 ----SESADPTAISGMKRRAHSDRVFDFPPCDPTMDLMIEAKEKEQAVFELC---
                                                                                                                                                                                                                                                                                 213 -SRSSAKNL-----EKESTMNI,DDHAPREMFD------CLDKPIPWRGRLGYACLN-
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   349 HRLTMHPGQYTQIASPRE-----VVVDSAIRDL------AYHDEIL-----
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             ---KDAV----
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    548 RRYELQNPPCPLEIMGPEYDQTRDGYYPPGAEK-----RLTARKRRSRKEEVEED
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   A; Experimental source: strain $288C
R; Koelling, R.; Nguyen, T.; Chen, E.Y.; Botstein, D.
Mol. Gen. Genet. 237, 359-369, 1993
A; Title: A new yeast gene with a myosin-like heptad repeat structure.
A; Reference number: $31207; MUID:93247549
                                                                                            170 K-----QKEEEEEYVEEVDEKS---LKNESSSDEFEPVVPEQLETPISKRRR----
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     311 MRVSSDLFPFASHAKYGY----TLEFAQSHLEEVGKL---
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             384 -----SRMKLNEOLN--
```

```
17;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           1566 MEQSGEIDVVLRKQLEARVQEKQ-----KELENEYNKKLQEELKDVPHSSHISDDERDK 1619
                                                                                                                                                                                                                                                                                                                                    A;Residues: 1-300,'A',302-1875 <KOE>
A;Cross-references: EMBL:L01992; NID:g171958; PIDN:AAA34783.1; PID:g171959
C;Genetics:
A;Gene: SGD:MLP1
A;Cross-references: SGD:S0001803; MIPS:YKR095w
A;Map position: 11R
                                                                                                                                                                                                        Query Match
4.6%; Score 144; DB 2; Length 1875;
Best Local Similarity 21.4%; Pred. No. 4.6;
Matches 92; Conservative 71; Mismatches 164; Indels 102; Gaps
                                                                                                                                                                                                                                                                                                                                                                                               148 UNYNSSSESE------SEEQISKATKKVKQKEEEEFYVEEVDEKSLKN 188
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    240 LDKPIPWRGRLGYACLNTILRSMKERVF----CSRTCRITTIQRDGLESVKQL----- 288
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          289 -----GTQN-VLDLIKLVEWNHNFGIHFMRVSSDLFPFASHAKYG----YTLEFAQS 335
                                                                                                                                                                                                                                                                                                   88 ETELANISGPHKKSTSTSTRKRARSSKKKATDSVSDKIDESVASYDSSTHLRRSSRSKKP 147
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        189 ESSSDEFEPVVPEQLETPISKRRSRSSAKNLEKESTMNLDD------HAPREMFDC 239
                                                                                                                                                                                                                                                                                                                                                                                                  δλ
                                                                                                                                                                                                                                                                                                                                                                                                                                            Ωp
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                QQ
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              qq
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 δy
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          δy
```

1677 ESPPKSVNNVQNPLLGLPRKIEENSNSPFNPLLSGEKLLKLNSKSSSGGFNPFTSPSPNK 1736

qq

qq

ò

οy

1795 DSSNRNVQSEEDT-----EKKKEGEPVKRGEAIEEQTKSNKRPIDEV-GELKNDE 1843

385 -RMKLNEQLNKDAVLIIHLGGTFEGKK-----ETLDRFRKNYQRLSDSVKARLVLEN 435

336 HLEEVG----KLANKYNHRLTMHPGQYTQIASPREVVVDSAIRDLAYHDEILS----- 384

Search completed: January 15, 2002, 13:52:22 Job time: 391 sec

436 DDVSWSVQD 444 || : : : 1844 DDTTENINE 1852

9FQ 1D

Tue Jan 15, 14:12:07 2002

NO 2 is an elowar

Copyright (c) 1993 - 2000 Compugen Ltd. ///LE C | LLV | 21/4

OM protein - protein search, using sw model

January 15, 2002, 13:50:21; Search time 134.53 Seconds Run on:

(without alignments) 900.271 Million cell updates/sec

4349 1 MTKLPILGYWKIKGLVQPTR.....KRLTARKRSRKEEVEEDEK 828 US-09-724-296-2 Title: Perfect score: Sequence:

BLOSUM62 Gapop 10.0 , Gapext 0.5 Scoring table:

473505 seqs, 146272329 residues Searched:

473505 Total number of hits satisfying chosen parameters:

Minimum DB seq length: 0 Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

SPTREMBL_17:* Database

sp_unclassified:* sp_human:*
sp_invertebrate:* sp_vertebrate:* sp_organelle:* sp_archea:* sp_bacteria:* sp_rodent:* sp_mammal:* sp_plant:* sp_virus:* sp_phage:* sp_fungi:* sp_mhc:* 10:

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

		Description	Q10988 schizosacch	009195 escherichia	009194 escherichia	009193 escherichia		009191 escherichia	Q94745 schistosoma	Q01408 neurospora	Q25595 clonorchis	Q9xyl9 fasciola gi	Q9n0v4 bos taurus	097117 boophilus m	O35660 mus musculu	Q9tsm5 macaca fasc	Q9beb0 macaca fusc	Q9tsm4 macaca fasc	Q27653 echinococcu	O16058 echinococcu	Q9wu21 rattus norv
	;	ID	Q10988	009195	009194	009193	009192	009191	094745	001408	025595	O9XYL9	Q9N0V4	097117	035660	Q9TSM5	Q9BEB0	Q9TSM4	027653	016058	Q9WU21
	:) H	٣	7	7	7	7	7	Ŋ	٣	ស	ง	9	Ŋ	11	9	9	9	Ŋ	S	11
	;	Match Length DB	599	262	268	245	243	244	218	929	218	218	218	223	218	218	221	218	219	219	218
æ	Query	Match	71.6	27.6	27.6	27.2	27.2	27.2	24.0	23.3	16.4	16.0	11.7	11.4	11.0	11.0	10.9	10.9	10.8	10.7	10.6
		Score	3112	1200.5	1200.5	1182	1181	1181	1043	1011.5	712	869	510.5	496.5	479	476.5	476	473.5	468.5	467.5	460.5
	Result	S S	7	2	٣	4	S	9	7	æ	6	10	11	12	13	14	15	16	17	18	19

Q9dce8 mus musculu Q9dd25 mus musculu Q60550 homo sapien Q9bea9 macaca fusc Q9d518 mus musculu	0921b2 rattus norv 005465 homo sapien 09mzb4 capra hircu 09u582 psoroptes o	Q9494e9 bacillus ha Q949p8 bacillus ha Q29583 sus scrofa Q02636 brugia mala Q9du5 brachydanio P81942 bufo bufo Q27711 onchocerca Q15690 homo scapien	O00460 homo sapien Q9nhb2 mytllus edu Q9tty8 capra hircu Q9s0m9 deinococcus Q9rte6 deinococcus Q9rte7 oncorpynchu Q9ue37 homo sapien Q9dcul mus musculu
Q9DCE8 Q9DD25 060550 Q9BEA9 09D5J8	Q921B2 Q05465 Q9M2B4 Q9U582 Q9L4F0	Q9L4E9 Q9K9P8 Q29583 Q0DDU5 P81942 Q27711	000460 09NHB2 09TTY8 09SOM9 09RTE6 09W647 09UE37
11 11 4 6	11 6 7	2 2 3 4 4 5 1 1 3 4	4 6 6 6 7 7 1 3 1 3 1 1 1 1 1 1 1 1 1 1 1 1 1 1
218 218 225 225 232	225 195 188 219 317	325 322 125 208 208 210	210 180 209 305 326 216 221
10.5 10.5 10.3 10.1	10.00 9.9 4.00 7.88	8 6 6 6 6 6 6 6 6 6 6 6 6 6 6 6 6 6 6 6	ი ი ი ი ი 4 4 ი 4 4 6 6 0 0 0 4
457.5 455.5 446.5 440.5	434.5 430.5 407.5 359	359 306.5 285.5 253.5 246 244.5 240	240 236 229.5 229.5 229.5 213 212.5 193.5
20 22 23 24	25 27 28 28 29	33 3 3 3 3 3 3 3 3 3 3 3 3 3 3 3 3 3 3	88844444 800113844
	3		

ALIGNMENTS

```
Takao M., Yonemasu R., Yamamoto K., Yasui A.;
"Characterization of a UV endonuclease gene from the fission yeast Schizosaccharomyces pombe and its bacterial homolog.";
Nucleic Acids Res. 24:1267-1271(1996).
                                                                                                                                                                                                                                                                                                                                                                                                                                                            Lyne M., Rajandream M.A., Barrell B.G., Lucas M., Gaillardin C.; Submitted (JUN-1998) to the EMBL/GenBank/DDBJ databases.
EMBL; D78571; BAA11415.1; -.
EMBL; U78487; AAC49664.1; -.
                                                                                                                                                                                                                                                                                                                                                            Davey S., Nass M.L., Ferrer J.V., Sidik K., Eisenberger A., Mitchell D.L., Freyer G.A.;
                                                                                                                                                                                                                                                                                                                                                                                        "The fission yeast UVDR DNA repair pathway is inducible."; Nucleic Acids Res. 25:1002-1008(1997).
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   C09ED7219D5895BE CRC64;
                                                                                                               UVDE OR UVEI+.
Schizosaccharomyces pombe (Fission yeast).
Bukaryota: Fungi; Ascōmycotā; Schizosaccharomycetales; Schizosaccharomycetales; Schizosaccharomyces.
NCBI_TaxID=4896;
                                                                     Last sequence update)
Last annotation update)
                            599 AA
                                                         Created)
                            PRT;
                                                                                                                                                                                                                                                                                                                                                   MEDLINE=97175806; PubMed=9023111;
                                                                                                                                                                                                                                  STRAIN=SP972;
MEDLINE=96188860; PubMed=8614629;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      68816 MW;
                                    010988; P87339;
01-NOV-1996 (TYEMBLEEL: 01, C.
01-NOV-1998 (TYEMBLEEL: 01, L.
UV-ENDONUCLEASE: 08, LA
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        EMBL; AL023859; CAA19577.1;
                            PRELIMINARY;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         Endonuclease.
SEQUENCE 599 AA;
                                                                                                                                                                                                                                                                                                                                                                                                                                       SEQUENCE FROM N.A.
                                                                                                                                                                                                                                                                                                                                     SEQUENCE FROM N.A.
                                                                                                                                                                                                                    SEQUENCE FROM N.A.
                                                                                                                                                                                                                                                                                                                                                                                                                                                      STRAIN-972H-;
                            010988
RESULT
Q10988
```

Query Match

71.6%; Score 3112; DB 3; Length 599;

```
ó
                                                  230 MERLIKRNIQISKRIVFTILKQKAFKGNHPCVPSVCTITYSRFHCLPDTLKSLLPMSSKT 289
                                                                                                                                                                                                                                                                                         290 TLSMLPQVNIGANSFSAETPVDLKKENETELANISGPHKKSTSTSTRKRARSSKKKATDS 349
                                                                                                                         350 VSDKIDESVASYDSSTHLRRSSRSKKPVNYNSSSESEEGISKATKKVKQKEEEEYVEE 409
                                                                                                                                                                                410 VDEKSLKNESSSDEFEPVVPEQLETPISKRRRSRSSAKNLEKESTMNLDDHAPREMFDCL 469
                                                                                                                                                                                                                                    470 DKPIPWRGRLGYACLNTILRSMKERVFCSRTCRITTIQRDGLESVKQLGTQNVLDLIKLV 529
                                                                                                                                                                                                                                                                                                                                 EWNHNFGIHFMRVSSDLFPFASHAKYGYTLEFAQSHLEEVGKLANKYNHRLTMHPGQYTQ 589
                                                                                                                                                                                                                                                                                                                                                590 IASPREVVVDSAIRDLAYHDEILSRMKLNEQLNKDAVLIIHLGGTFEGKKETLDRFRKNY 649
                                                                                                                                                                                                                                                                                                                                                                                                                  361 IASPREVVVDSAIRDLAYHDEILSRMKLNEQLNKDAVLIIHLGGTFEGKKETLDRFRKNY 420
                                                                                                                                                                                                                                                                                                                                                                                                                                             650 QRLSDSVKARLVLENDDVSWSVQDLLPLCQELNIPLVLDWHHHNIVPGTLREGSLDLMPL 709
                                                                                                                                                                                                                                                                                                                                                                                                                                                             710 IPTIRETWTRKGITQKQHYSESADPTAISGMKRRAHSDRVFDFPPCDPTMDLMIEAKEKE 769
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   0; Gaps
                                                                               1 MLRILKRNIQISKRIVFTILKQKAFKGNHPCVPSVCTITYSRFHCLPDTLKSLLPMSSKT 60
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           770 QAVFELCRRYELQNPPCPLEIMGPEYDQTRDGYYPPGAEKRLTARKRRSRKEEVEEDEK 828
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        Bacteria; Proteobacteria; gamma subdivision; Enterobacteriaceae;
                            Indels
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            Cahill M.A.;
Submitted (JAN-1997) to the EMBL/GenBank/DDBJ databases.
EMBL; U84572; AAB41883.1; · HSSP; P08515; IGNE.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       Hipskind R.H., Delaney P.;
Submitted (FEB-1997) to the EMBL/GenBank/DDBJ databases
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               30392 MW; BA4C4593CB20226E CRC64;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 01-JUL-1997 (TrEMBLrel. 04, Created)
01-JUL-1997 (TrEMBLrel. 04, Last sequence update)
01-JUN-2001 (TrEMBLrel. 17, Last annotation update)
Pred. No. 6e-190;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           262 AA.
                          0; Mismatches
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           PRT;
          100.08;
                          Matches 599; Conservative
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       Escherichia coli.
Plasmid pGEX-2T-6His-PL2.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   InterPro; IPR000521; GST.
Pfam; PF00043; GST; 1.
SEQUENCE 262 AA; 30392
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           PRELIMINARY;
           Best Local Similarity
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               SEQUENCE FROM N.A.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          SEQUENCE FROM N.A.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               NCBI_TaxID=562;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                Escherichia
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              GST-6HIS.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        009195
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        009195
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               ~
                                                                                                                                                                                                                                                                                                                                 530
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               RESULT
009195
                                                     ò
                                                                              d
                                                                                                          ô
                                                                                                                                     g
                                                                                                                                                                 δ
                                                                                                                                                                                          qq
                                                                                                                                                                                                                      δŏ
                                                                                                                                                                                                                                               g
                                                                                                                                                                                                                                                                            QY
                                                                                                                                                                                                                                                                                                     q
                                                                                                                                                                                                                                                                                                                                 ò
                                                                                                                                                                                                                                                                                                                                                          g
                                                                                                                                                                                                                                                                                                                                                                                      ŏ
                                                                                                                                                                                                                                                                                                                                                                                                                  pp
                                                                                                                                                                                                                                                                                                                                                                                                                                             ŏ
                                                                                                                                                                                                                                                                                                                                                                                                                                                                        op
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    δŏ
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             qq
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           δ
```

```
1;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           ;
                                                                                                                                            63 VKLTQSMAIIRYIADKHNMLGGCPKERAEISMLEGAVLDIRYGVSRIAYSKDFETLKVDF 122
                                                                                                                                                                                                      65 VKLTQSMAliryladkhnmlggcpkeraeismlegavldirygvsriayskdfetlkvdf 124
                                                                                                                                                                                     125 LSKLPEMLKMFEDRLCHKTYLNGDHVTHPDFMLYDALDVVLYMDPMCLDAFPKLVCFKKR 184
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                123 LSKLPEMLKMFEDRLCHKTYLNGDHVTHPDFMLYDALDVVLYMDPMCLDAFPKLVCFKKR 182
                                  1; Gaps
                                                                                    3 PILGYWKIKGLVQPTRLLLEYLEEKYEEHLYERDEGDKWRNKKFELGLEFPNLPYYIDGD 62
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 65 VKLTQSMATTRYLADKHNMLGGCPKERAETSMLEGAVLDIRYGVSRIAYSKDFETLKVDF 124
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            125 LSKLPEMLXMFEDRLCHKTYLNGDHVTHPDFMLYDALDVVLYMDPMCLDAFPKLVCFKKR 184
                                                             5 PILGYWKIKGLVQPTRLLLEYLEEKYEEHLYERDEGDKWRNKKFELGLEFPNLPYYIDGD 64
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       1; Gaps
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      3 PILGYWKIKGLVOPTRLLLEYLEEKYEEHLYERDEGDKWRNKKFELGLEFPULPYYIDGD 62
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    5 PILGYWKIKGLVQPTRLLLEYLEEKYEEHLYERDEGDKWRNKKFELGLEFPNLPYYIDGD 64
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    Bacteria; Proteobacteria; gamma subdivision; Enterobacteriaceae;
 Score 1200.5; DB 2; Length 262;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        27.6%; Score 1200.5; DB 2; Length 268; 97.8%; Pred. No. 7.4e-69;
                                                                                                                                                                                                                                                 185 IEAIPQIDKYLKSSKYIAWPLQGWQATFGGGDHPPKSDHLVPRGSMLRL 233
                                                                                                                                                                                                                                                                   Indels
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          185 IEAIPQIDKYLKSSKYIAWPLQGWQATFGGGDHPPKSDHLVPRGSMLRL 233
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            183 IEAIPQIDKYLKSSKYIAWPLQGWQATFGGGDHPPKSD-LVPRGSPVEM 230
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       Indels
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            Submitted (JAN-1997) to the EMBL/GenBank/DDBJ databases. EMBL, U84571; AAB41882.1; -. HSSP; P08515; IGNE.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          31000 MW; 85820375FE60E625 CRC64;
                                                                                                                                                                                                                                                                                                                                                                                                         Last sequence update)
Last annotation update)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      2;
                Pred. No. 7.2e-69;
2; Mismatches 2;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 01-JUL-1997 (TrEMBLrel. 04, Created)
01-JUL-1997 (TrEMBLrel. 04, Last sequence update)
                                                                                                                                                                                                                                                                                                                                                              268 AA.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       245 AA
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       2; Mismatches
                                                                                                                                                                                                                                                                                                                                                                                          Created)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       PRT;
27.6%; 97.8%;
                                                                                                                                                                                                                                                                                                                                                                                       01-JUL-1997 (TrEMBLrel. 04, 01-JUL-1997 (TrEMBLrel. 04, 01-JUN-2001 (TrEMBLrel. 17,
Query Match 27.69
Best Local Similarity 97.89
Matches 224; Conservative
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            IPR000521; GST.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    Matches 224; Conservative
                                                                                                                                                                                                                                                                                                                                                             PRELIMINARY;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       PRELIMINARY;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    Pfam; PF00043; GST; 1 SHOUENCE 268 AA; 3:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     Plasmid pGEX-2T-6H.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       Best Local Similarity
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              SEQUENCE FROM N.A.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                     Escherichia coli.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  NCBI_TaxID=562;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                Cahill M.A.;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     Escherichia
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              InterPro;
                                                                                                                                                                                                                                                                                                                                                                                                                                         GST-6HIS.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        Ouery Match
                                                                                                                                                                                                                                                                                                                                                                          009194;
                                                                                                                                                                                                                                                                                                                                                         009194
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     009193
                                                                                                                                                                                                                                                                                                                             e
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       4
                                                                                                                                                                                                                                                                               183
                                                                                                                                                                                                                                                                                                                             RESULT
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     009193
                                                              δλ
                                                                                          pp
                                                                                                                          Óλ
                                                                                                                                                      q
                                                                                                                                                                                     δŏ
                                                                                                                                                                                                                  qq
                                                                                                                                                                                                                                                 Óγ
                                                                                                                                                                                                                                                                                                                                                           ID
                                                                                                                                                                                                                                                                                                                                                                          qq
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                qq
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     AC AC DT
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 QΥ
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               qd
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          a
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              ŏ
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           9
```

```
SEQUENCE FROM N.A.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              SEQUENCE FROM N.A.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               Plasmid pGEX-6P-1
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    Escherichia coli.
     NCBI_TaxID=562;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          NCBI_TaxID=562;
                                                                                                                          Colonno R.J.;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 Escherichia.
                                                                                                                                                                                                                                                             SEQUENCE
                                                                                                                                                                                                                                                                                           Query Match
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        009191;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             009191
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          RESULT
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     009191
     q
                                                                                                                                                                                                                                                                                                                                   ò
                                                                                                                                                                                                                                                                                                                                                                                           qq
                                                                                                                                                                                                                                                                                                                                                                                                                                    qq
                                                                                                                                                                                                                                                                                                                                                                          õ
                                                                                                                                                                                                                                                                                                                                                                                                                ŏ
                                                                                                                                                                                                                                                                                                                                                                                                                                                        δ
                                                                                                                                                                                                                                                                                                                                                                                                                                                                          Д
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  7
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       65 VKLTQSMAIIRYIADKHNMLGGCPKERAEISMLEGAVLDIRYGVSRIAYSKDFETLKVDF 124
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                LSKLPEMLKMFEDRLCHKTYLNGDHVTHPDFMLYDALDVVLYMDPMCLDAFPKLVCFKKR 184
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            123 LSKLPEMLKMFEDRLCHKTYLNGDHVTHPDFMLYDALDVVLYMDPMCLDAFPKLVCFKKR 182
                                                                                                                                                                                                                                                                                                                                                                                 8; Gaps
                                                                                                                                                                                                                                                                                                                                                                                                             5 PILGYWKIKGLVQPTRLILEYLEEKYEEHLYERDEGDKWRNKKFELGLEFPNLPYYIDGD 64
                                                                                                                                                                                                     "Substrate requirements of human rhinovirus 3C protease for peptide
                                                                                                                      "Single-step purification of polypeptides expressed in Escherichia coli as fusions with glutathione S-transferase."; Gene 67:31-40(1988).
                                                  Bacteria; Proteobacteria; gamma subdivision; Enterobacteriaceae;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    Plasmid pGEX-6P-2.
Bacteria; Proteobacteria; gamma subdivision; Enterobacteriaceae;
                                                                                                                                                                                                                                                                                                                                                          27.2%; Score 1182; DB 2; Length 245; 92.2%; Pred. No. 9.8e-68; tive 1; Mismatches 10; Indels
                                                                                                                                                            SEQUENCE OF 221-227 FROM N.A.
MEDLINE-90264387; PubMed-2160953;
Cordingley M.G., Callahan P.L., Sardana V.V., Garsky V.M.,
                                                                                                                                                                                                                                                    Dunst R.W., Bell P.A., English J.D.;
Submitted (NOV-1996) to the EMBL/GenBank/DDBJ databases
EMBL; U78874; AAB37352.1;
HSSP: P08515; 10NE.
                                                                                                                                                                                                                                                                                                                             245 AA; 28327 MW; D51470E8A48C6CC2 CRC64;
01-JUN-2001 (TrEMBLrel. 17, Last annotation update) GLUTATHIONE S-TRANSFERASE.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     Last sequence update)
Last annotation update)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       243 AA
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           Created)
                                                                                                                                                                                                                 cleavage in vitro.";
J. Biol. Chem. 265:9062-9065(1990).
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       PRT;
                                                                                                   MEDLINE=88329742; PubMed=3047011;
Smith D.B., Johnson K.S.;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       01-JUL-1997 (TrEMBLrel. 04, 01-JUL-1997 (TrEMBLrel. 04, 01-JUN-2001 (TrEMBLrel. 17, GLUTATHIONE S-TRANSFERASE.
                                                                                                                                                                                                                                                                                              InterPro; IPR000521; GST.
Pfam; PF00043; GST; 1.
Transferase; Plasmid.
                                                                                                                                                                                                                                                                                                                                                                    Best Local Similarity 92.28
Matches 226; Conservative
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      PRELIMINARY;
                                                                                         SEQUENCE FROM N.A.
                                                                                                                                                                                                                                             SEQUENCE FROM N.A.
                                        Plasmid pGEX-6P-3
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            Escherichia coli.
                              Escherichia coli
                                                                     NCBI_TaxID=562;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                241 SKRIV 245
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    239 SGRIV 243
                                                                                                                                                                                              Colonno R.J.;
                                                         Escherichia
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        Escherichia.
                                                                                                                                                                                                                                                                                                                             SEQUENCE
                                                                                                                                                                                                                                                                                                                                                           Query Match
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               009192;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     009192
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  Ŋ
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 125
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  RESULT
009192
q
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   qq
                                                                                                                                                                                                                                                                                                                                                                                                                     g
                                                                                                                                                                                                                                                                                                                                                                                                                                         ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                ŏ
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                ŏ
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       OC OC OC OC
```

```
65 VKLTQSMAIIRYIADKHNMLGGCPKERAEISMLEGAVLDIRYGVSRIAYSKDFETLKVDF 124
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              125 LSKLPEMLKMFEDRLCHKTYLNGDHVTHPDFMLYDALDVVLYMDPMCLDAFPKLVCFKKR 184
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    Gaps
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          3 PILGYWKIKGLVQPTRLLLEYLEEKYEEHLYERDEGDKWRNKKFELGLEFPULPYYIDGD 62
                                                                                                                                                                                                                                                                                            "Substrate requirements of human rhinovirus 3C protease for peptide cleavage in vitro."; J. Biol. Chem. 265:9062-9065(1990).
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           5 PILGYWKIKGLVQPTRLLLEYLEEKYEEHLYERDEGDKWRNKKFELGLEFPNLPYYIDGD 64
                                                                              Single-step purification of polypeptides expressed in Escherichia
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 "Single-step purification of polypeptides expressed in Escherichia coli as fusions with glutathione S-transferase."; Gene 67:31-40(1988).
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 4
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 Bacteria; Proteobacteria; gamma subdivision; Enterobacteriaceae;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            Length 243;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 185 IEAIPQIDKYLKSSKYIAWPLQGWQATFGGGDHPPKSDHLV----PRGS 229
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             Cordingley M.G., Callahan P.L., Sardana V.V., Garsky V.M.,
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              3; Indels
                                                                                                                                                                                                                                                                                                                                                                                                                                Dunst R.W., Bell P.A., English J.D.;
Submitted (NOV-1996) to the EMBL/GenBank/DDBJ databases.
EMBL: U78873; AAB37349.1; -.
HSSP: P08515; IONE.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     243 AA; 27908 MW; 0003C434CD76C3A6 CRC64;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       Last sequence update)
Last annotation update)
                                                                                                          coli as fusions with glutathione S-transferase.";
Gene 67:31-40(1988).
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      27.2%; Score 1181; DB 2; 96.9%; Pred. No. 1.1e-67;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           0; Mismatches
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         01-JUL-1997 (TrEMBLrel. 04, Created)
01-JUL-1997 (TrEMBLrel. 04, Last segu
01-JUN-2001 (TrEMBLrel. 17, Last anno
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               PRT;
SEQUENCE FROM N.A. MEDLINE=88329742; PubMed=3047011;
                                                                                                                                                                                     SEQUENCE OF 221-227 FROM N.A. MEDLINE=90264387; PubMed=2160953;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               [2]
SEQUENCE OF 221-227 FROM N.A.
MEDLINE=90264387; PubMed=2160953;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  MEDLINE=88329742; PubMed=3047011;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          GLUTATHIONE S-TRANSFERASE.
                                                     Smith D.B., Johnson K.S.;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       InterPro; IPR000521; GST.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          Best Local Similarity 96.9
Matches 222; Conservative
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           Smith D.B., Johnson K.S.;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         PRELIMINARY;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                Pfam; PF00043; GST; 1.
Transferase; Plasmid.
```

ij

RA RA RL RL RL RL DR DR CDR SO SO

q δλ g δ qq ò QQ

```
3 PILGYWKIKGLVQPTRILLEYLGEEYEERLYERNEGDAWRNEKFKLGLEFPNLPYYIDGD 62
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         Matches 241; Conservative
                                                                                                                                                                                                                                                                                                                                                                                                                                                STRAIN=74-OR23-1A;
                                                                                                                                                                                                                                                                                                                                                             Neurospora crassa.
                                                                                                                                                                                                                                                                                                                                                                                                                                   SEQUENCE FROM N.A.
                                                                                                                                                                                                                                                                                                                              UV-ENDONUCLEASE.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      Endonuclease
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    SEQUENCE
                                                                                                                                                                                                                                                                       001408;
                                                                                                                                                                                                                                                      001408
                                                                                                                                                                                                                          æ
                                                                                                                                                                                                                             RESULT
                                                                                                                                                                                                                                          001408
                                                                                                                                                                                                                                                                       Öζ
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               ŏγ
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           qq
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         δ
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     Dp
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  ŏ
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             qq
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           ŏ
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       QQ
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   δ
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                qq
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              οy
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     q
                                                                                                                      qq
                                                                                                                                                                              qq
                                                            qq
                                                                                                                                                    Qγ
                                ò
                                                                                            ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  .
0
                                                                                                                                                                                                                                                                                     ij
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    Grams S.V., Grams R., Korge G., Viyanant V., Upatham S.; "Clonning and sequence analysis of the 26 kDa glutathiones-transferase gene of Schistosoma mekongi.";
                                                                                                                                                                                                                                                                                                                                                                                           63 VKLTQSMAIIRYIADKHNMLGGCPKERAEISMLEGAVLDIRYGVSRIAYSKDFETLKVDF 122
                                                                                                                                                                                                                                                                                                                                                                                                                                                    123 LSKLPEMLKMFEDRLCHKTYLNGDHVTHPDFMLYDALDVVLYMDPMCLDAFPKLVCFKKR 182
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     Eukaryota; Metazoá; Platyhelminthes; Rhabditophora; Neodermata;
Trematoda; Digenea; Strigeidida; Schistosomatoldea; Schistosomatidae;
                                                                                                                                                                                                                                                                                                                                                                           65 VKLTQSMAIIRYIADKHNMLGGCPKERAEISMLEGAVLDIRYGVSRIAYSKDFETLKVDF 124
                                                                                                                                                                                                                                                                                                                                                                                                                                     125 LSKLPEMLKMFEDRLCHKTYLNGDHVTHPDFMLYDALDVVLYMDPMCLDAFPKLVCFKKR 184
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  0; Gaps
                                                                                                                                                                                                                                                                                     4; Gaps
                                                                                                                                                                                                                                                                                                                                    3 PILGYWKIKGLVQPTRLLLEYLEEKYEEHLYERDEGDKWRNKKFELGLEFPNLPYYIDGD 62
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          5 PILGYWKIKGLVQPTRLLLEYLEEKYEEHLYERDEGDKWRNKKFELGLEFPNLPYYIDGD 64
                                                                                                                                                                                                                                                                                                                   5 PILGYWKIKGLVQPTRLLLEYLEEKYEEHLYERDEGDKWRNKKFELGLEFPNLPYYIDGD 64
                                "Substrate requirements of human rhinovirus 3C protease for peptide cleavage in vitro.";
J. Biol. Chem. 265:9062-9065(1990).
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       Southeast Asian J. Trop. Med. Public Health 28:570-574(1997).
EMBL: Y07663; CAA68944.1; -- HSSP; P08515; 1GTA.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    24.0%; Score 1043; DB 5; Length 218; 89.8%; Pred. No. 5.8e-59;
                                                                                                                                                                                                                                                      Score 1181; DB 2; Length 244; Pred. No. 1.1e-67;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             Cordingley M.G., Callahan P.L., Sardana V.V., Garsky V.M.,
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  12; Indels
                                                                                                                                                                                                                                                                                     3; Indels
                                                                                                  Dunst R.W., Bell P.A., English J.D.;
Submitted (NOV-1996) to the EMBL/GenBank/DDBJ databases.
EMBL; U78872; AAB37346.1;
HSSP; P08515; IGNE.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          218 AA; 25486 MW; D3080620B19DE23A CRC64;
                                                                                                                                                                                                            244 AA; 28430 MW; 7461B31F17027042 CRC64;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             Last sequence update)
Last annotation update)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  10; Mismatches
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  218 AA
                                                                                                                                                                                                                                                                                       0; Mismatches
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  PRT;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             01-FEB-1997 (TrEMBLrel. 02, Created)
01-FEB-1997 (TrEMBLrel. 02, Last sequ
01-JUN-2001 (TrEMBLrel. 17, Last anno
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              STRAIN-THAILAND;
MEDLINE-98222395; Pubmed-9561610;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          KDA GLUTATHIONE S-TRANSFERASE.
                                                                                                                                                                                                                                                           27.2%;
96.9%;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         InterPro; IPR000521; GST.
Pfam; PF00043; GST; 1.
Transferase.
SEQUENCE 218 AA; 25486
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                Matches 194; Conservative
                                                                                                                                                                  InterPro; IPR000521; GST.
                                                                                                                                                                                                                                                                                       Matches 222; Conservative
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    PRELIMINARY;
                                                                                                                                                                                Pfam; PF00043; GST; 1.
Transferase; Plasmid.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   Similarity
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          Schistosoma mekonqi.
                                                                                                                                                                                                                                                                         Local Similarity
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             SEQUENCE FROM N.A.
                                                                                             SEQUENCE FROM N.A.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                NCBI_TaxID=38744;
                          Colonno R.J.;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     Query Match
Best Local S
                                                                                                                                                                                                                 SEQUENCE
                                                                                                                                                                                                                                                           Query Match
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                094745;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  094745
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        RESULT
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         094745
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    δŏ
```

```
339 ARSSKKKATDSVSDKIDESVASYDSST----HLRRSSR-----SKKPVNYNSSSESE- 386
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       234 SIVDHRHPLQFEDEPEHHLKNKPDKSKEPQDELGHKFVQELGLANARDIVKMLCWNEKYG 293
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              537 IHFMRVSSDLFPFASHAKYGYTL-EFAQSHLEEVGKLANKYNHRLTMHPGQYTQIASPRE 595
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                294 1RFLRLSSEMPPFASHPVHGYKLAPFASEVLAEAGRVAAELGHRLTTHPGQFTQLGSPRK 353
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              596 VVVDSAIRDEAYHDEILSRMKINEQLNKDAVLIIHLGGTFEGKKETLDRFRKNYQRLSDS 655
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 387 -----SEQISKATKKVKQKEEEEYVEEVDEK 413
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              63 AGRAMGKIJANGHCLREGKEQEEGVKMAIEGLARMERRLQRATKROKKQLEED----- 114
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                414 SLKNESSSDEFEPVVPEQLETPISKRRS-----RSSAKNLEKESTMNLDDHA----- 461
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              115 GIPVPSVVSRF-PTAPYHHKSTNAEEREAKEPVLKTHSKDVEREAEIGVDDVVKMEPAAT 173
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                462 ----PREMFDCLDK------PIPWRGRLGYACLNTILRSMKERVFCSRTCRIT 504
                                                                                                                                                            "A eukaryotic gene encoding an endonuclease that specifically repairs
                           63 VKLTQSMAITRYIADKHNMLGGSGKEKAELTMLEGAVSDIRSGVSRIAYNKDFETLKVDF 122
65 VKLTQSMAIIRYIADKHNMLGGCPKERAEISMLEGAVLDIRYGVSRIAYSKDFETLKVDF 124
                                                                                                                         125 LSKLPEMLKMFEDRLCHKTYLNGDHVTHPDFMLYDALDVVLYMDPMCLDAFPKLVCFKKR 184
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        Indels 181; Gaps
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 3 SRKSKAAALDTPQSESSTFSSTLDSSAPSPARNLRRSGRNILQPSSEKDRDHEKRSGEEL 62
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          Yajima H., Takao M., Yasuhira S., Zhao J.H., Ishii C., Inoue H.,
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     Eukaryota; Fungi; Ascomycota; Pezizomycotina; Sordariomycetes; Sordariales; Sordariaceae; Neurospora.
NCBL_TaxID=5141;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  Query Match 23.3%; Score 1011.5; DB 3; Length 656; Best Local Similarity 36.6%; Pred. No. 2.6e-56;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        656 AA; 74413 MW; 4B231E7A8735C86F CRC64;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  01-NOV-1996 (TrEMBLrel. 01, Created)
01-NAY-1999 (TrEMBLrel. 10, Last sequence update)
01-NAY-1999 (TrEMBLrel. 10, Last annotation update)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            ; Pred. No. 2.6e-56; 88; Mismatches 149;
                                                                                                                                                                                                                                                             185 IEAIPQIDKYLKSSKYIAWPLQGWQATFGGGDHPPK 220
                                                                                                                                                                                                                                                                                                   656 AA.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        DNA damaged by ultraviolet light.";
EMBO J. 14:2393-2399(1995).
EMBL; D11392; CAB20113.1;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           MEDLINE=95292980; PubMed=7774597;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  PRELIMINARY; PRT;
```

18;

PRELIMINARY;

```
;
0
                                                                        533 VFELMRIFKLPGFEKINDMVPYDRDDENRPAPPVKAPKKKGGKRKRTTDEEAAEPEEVD 592
                                                                                                                                                                                                                                                                                                                                                   Clonorchis sinensis.

Eukaryota; Metazoa; Platyhelminthes; Rhabditophora; Neodermata;
Trematoda; Digenea; Opisthorchiida; Opisthorchiata; Opisthorchiidea;
Opisthorchiidae; Clonorchis.
656 VKARLVLENDDVSWSVQDLLPLCQELNIPLVLDWHHHNIV--PGTLREGSLDLM--PLIP 711
                                                      712 TIRETWTRKGITQKQHYSESADPTAISGMKRRAHSDRVFDFPPCDPTMDLMIEAKEKEQA 771
                                                                                                             Hong S.-J., Lee D.-H.;
"Cloning and over expression of 26 kDa glutathione S-transferase from
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             Thesis (1996), Parasitology, Chung-Ang University College of Medicine.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         65 VKLTQSMAIIRYIADKHNMLGGCPKERAEISMLEGAVLDIRYGVSRIAYSKDFFTLKVDF 124
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           LSKLPEMLKMFEDRLCHKTYLNGDHVTHPDFMLYDALDVVLYMDPMCLDAFPKLVCFKKR 184
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    795 --YDQTRDG-----EEVEED 826
                                                                                                                                                                                         593 TAADDVKDAPEGPKEVPEEERAMGGPYNRVYWPLGCEEWLKPKKREVKKGKVPEEVEDE 651
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              0; Gaps
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   5 PILGYWKIKGLVQPTRLLLEYLEEKYEEHLYERDEGDKWRNKKFELGLEFPNLPYYIDGD 64
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      Length 218;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            52; Indels
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  Submitted (APR-1999) to the EMBL/GenBank/DDBJ databases
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     218 AA; 25038 MW; 7CB17C7B837A0B7C CRC64;
                                                                                                                                                                                                                                                                                           01-NOV-1996 (TrEMBLrel. 01, Created)
01-NOV-1999 (TrEMBLrel. 12, Last sequence update)
01-JUN-2001 (TrEMBLrel. 17, Last annotation update)
PUTATIVE GLUTATHIONE TRANSFERASE.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                16.4%; Score 712; DB 5; 59.3%; Pred. No. 6.5e-38;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          185 IEAIPQIDKYLKSSKYIAWPLQGWQATFGGGDHPPK 220
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         Conservative 36; Mismatches
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   EMBL; L47992; AAB46369.3;
HSSP; P31670; 1FHE.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          InterPro; IPR000521; GST.
Pfam; PF00043; GST; 1.
                                                                                                                                                                                                                                                                    PRELIMINARY;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 Clonorchis sinensis.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            Similarity
                                                                                                                                                                                                                                                                                                                                                                                                                                       SEQUENCE FROM N.A.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        SEQUENCE FROM N.A.
                                                                                                                                                                                                                                                                                                                                                                                                           NCBI_TaxID=79923;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        Transferase,
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         Matches 128;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    SEQUENCE
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               Query Match
Best Local S
                                                                                                                                                                                                                                                                   025595
                                                                                                                                                                                                                                          σ
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   125
                                                                                                                                                                                                                                          RESULT
                                                                                g
                                                                                                             ò
                                                                                                                                         qq
                                                                                                                                                                                             q
                                                                                                                                                                                                                                                                      ò
                          셤
                                                      ò
                                                                                                                                                                    ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               qq
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          ò
```

10

RESULT 09XYL9

```
67 LTQSMAIIRYIADKHNMLGGCPKERAEISMLEGAVLDIRYGVSRIAYSKDFETLKVDFLS 126
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      127 KLPEMLKMFEDRLCHKTYLNGDHVTHPDFMLYDALDVVLYMDPMCLDAFPKLVCFKKRIE 186
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       7 LGYWKIKGLVQPTRLLLEYLEEKYEEHLYERDEGDKWRNKKFELGLEFPNLPYYIDGDVK 66
                                                                                                                                                                                                                                                                                                                                                                                                                                                     STRAIN=THAILAND;
Grams S.V. Grams R., Sobhon P., Viyanant V., Upatham E.S.;
"Molecular cloning of expressed antigens from Fasciola gigantica.";
Submitted (DEC-1998) to the EMBL/GenBank/DDBJ databases.
EMBL; AF112567; AAD23997.1; '.
HSSP; P31670; IFHE.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                Mammalia; Eutheria; Cetartíodactyla; Ruminantia; Pecora; Bovoidea;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  Bos taurus (Bovine).
Eukaryota, Metazoa, Chordata, Craniata, Vertebrata, Euteleostomi,
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   "A lens glutathione S-transferase, class mu, with thiol-specific
                                                                                                               Eukaryota; Metazoa; Platyhelminthes; Rhabditophora; Neodermata;
Trematoda; Digenea; Echinostomida; Echinostomata; Fascioloidea;
Fasciolidae; Fasciola.
                                                                                                                                                                                                                                                                                                                                                                                                            0;
                                                                                                                                                                                                                                                                                                                                                                               Length 218;
                                                                                                                                                                                                                                                                                                                                                                                                          53; Indels
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            antioxidant activity.";
Submitted (MAR-2000) to the EMBL/GenBank/DDBJ databases.
EMBL, AF249588; AAF64308.1;
InterPro; IPR000521; GST.
InterPro; IPR00321; GST.
                                                                                                                                                                                                                                                                                                                                 218 AA; 25337 MW; E9423D75C3F2EEAF CRC64;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        01-007-2000 (TrEMBLrel. 15, Created)
01-0CT-2000 (TrEMBLrel. 15, Last sequence update)
01-JUN-2001 (TrEMBLrel. 17, Last annotation update)
                                          Last sequence update)
Last annotation update)
                                                                                                                                                                                                                                                                                                                                                                             Score 698; DB 5;
Pred. No. 5.1e-37;
                                                                                                                                                                                                                                                                                                                                                                                                         35; Mismatches
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 185 DLPKIKAYMESEKFIKWPLNSWTASFGGGDAAP 217
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             187 AIPQIDKYLKSSKYIAWPLQGWQATFGGGDHPP 219
                              Created)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   PRT;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        CLASS MU GLUTATHIONE S-TRANSFERASE.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     Jimenez-Asensio J.V., Garland D.;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      PRINTS; PR01267; GSTRNSFRASEM.
                         (TrEMBLrel. 12, C
(TrEMBLrel. 12, I
(TrEMBLrel. 17, I
                                                                                                                                                                                                                                                                                                                                                                             16.0%;
58.7%;
                                                                      GLUTATHIONE S-TRANSFERASE.
                                                                                                                                                                                                                                                                                                                                                                                                         Matches 125; Conservative
                                                                                                                                                                                                                                                                                        InterPro; IPR000521; GST.
Pfam; PF00043; GST; 1.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   PRELIMINARY;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            Bovidae; Bovinae; Bos.
NCBL_TaxID=9913;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       Pfam; PF00043; GST; 1.
                                                                                                 Fasciola gigantica
                                                                                                                                                                                                                                                                                                                                                                                           Best Local Similarity
                                                                                                                                                                                     SEQUENCE FROM N.A.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       SEQUENCE FROM N.A.
                                                                                                                                                         NCBI_TaxID=46835;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        TISSUE=EYE LENS;
                         01-NOV-1999
01-NOV-1999
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         01-OCT-2000
01-JUN-2001
                                                        01-JUN-2001
                                                                                                                                                                                                                                                                                                                  Transferase
                                                                                                                                                                                                                                                                                                                                 SEQUENCE
                                                                                                                                                                                                                                                                                                                                                                             Query Match
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              Q9N0V4;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                09N0V4
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      RESULT 11
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     Q9N0V4
ŏ
                                                                                                                                                                                                                                                                                                                                                                                                                                                                 q
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  q
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              Οy
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        qq
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     δλ
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               δλ
```

ö

```
035660
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               RESULT 14
                                                                                                                                                                                                                                                                                                                                                                                                                                                                         HSSE:
                               RESULT 13
                                                                                                                                                                                                                                                                                                  dene
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             O9TSM5
                                             032660
                                                                                           q
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  δy
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            q
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           qq
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     ŏλ
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           pp
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         AC AC
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                ..
--i
                                                                                  ;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           "Characterization and molecular cloning of a glutathione S-transferase from larvae of the cattle tick Boophilus microplus (Acari:
                                                                                                                                                                                                                          60 YIDGDVKLTQSMAIIRYIADKHNMLGGCPKERAEISMLEGAVLDIRYGVSRIAYSKDFET 119
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       120 LKVDFLSKLPEMLKMFEDRLCHKTYLNGDHVTHPDFMLYDALDVVLYMDPMCLDAFPKLV 179
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  123 LKGDYLKNLPASLKAFSDYLGTHKFFAGDNLTYVDFIAYEMLAQHLIFAPDCLKDFANLK 182
                                                                                                                                                         61 IDGDVKLTQSMAIIRYIADKHNMLGGCPKERAEISMLEGAVLDIRYGVSRIAYSKDFETL 120
                                                                                                                                                                                 64 IDGTHKLTQSNAILRYIARKHNLCGETEEEMIRVDILENQVMDVRLAMARICYSPDFEKL 123
                                                                                                                                                                                                            121 KVDFLSKLPEMLKMFEDRLCHKTYLNGDHVTHPDFMLYDALDVVLYMDPMCLDAFPKLVC 180
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         5; Gaps
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        5 PILGYWKIKGLVQPTRLLLEYLEEKYEEHLYERD-----EGDKWRNKKFELGLEFPNLPY 59
                                                                                  Gaps
                                                                                                          6 ILGYWKIKGLVQPTRLLLEYLEEKYEEHLYERDEG-----DKWRNKKFELGLEFPNLPYY 60
                                                                                                                                4 ILGYWDIRGLAHAIRLLLEYTDTNYEERQYSVGDAPDYDRSQWLNEKFKLGLDFPNLPYL 63
                                                                                                                                                                                                                                                                                                                                                                                                                               Eukāryota; Metažoa; Arthropoda; Chelicerata; Arachnida; Acari; Parasitiformes; Ixodida; Ixodidae; Boophilus.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     11.4%; Score 496.5; DB 5; Length 223; 47.2%; Pred. No. 3.4e-24; tive 28; Mismatches 80; Indels 5;
                                                         DB 6; Length 218;
                                                       11.7%; Score 510.5; DB 6; Length 347.8%; Pred. No. 4.2e-25; tive 35; Mismatches 67; Indels
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 Submitted (JUL-1998) to the EMBL/GenBank/DDBJ databases.
EMBL, AF077609; AAD15991.1; -.
HSSP; P20136; 16SU.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    223 AA; 25604 MW; F276ACE2D8259174 CRC64;
                   25635 MW; 3D02EA0F43C07B0A CRC64;
                                                                                                                                                                                                                                                                                                                                                                         01-MAY-1999 (TrEMBLrel. 10, Created)
01-MAY-1999 (TrEMBLrel. 10, Last sequence update)
01-JUN-2001 (TrEMBLrel. 17, Last annotation update)
GLUTATHIONE S-TRANSFERASE.
                                                                                                                                                                                                                                                                                                                                                   223 AA.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           180 CFKKRIEAIPQIDKYLKSSKYIAWPLQGWQATFG 213
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        PRT;
                                                                                                                                                                                                                                                             181 FKKRIEAIPQIDKYLKSSKYIAWPL 205
                                                                                                                                                                                                                                                                             184 FISRFEGLKKISAYMKSSRFLPGPL 208
                                                                                                                                                                                                                                                                                                                                                                                                                            Boophilus microplus (Cattle tick).
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           PRINTS; PR01267; GSTRNSFRASEM.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       InterPro; IPR000521; GST.
InterPro; IPR003081; GST_mu.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 Matches 101; Conservative
                                                                      Best Local Similarity 47.89
Matches 98; Conservative
                                                                                                                                                                                                                                                                                                                                                    PRELIMINARY;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                GST;
Transferase.
Transferase.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     Similarity
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        SEQUENCE FROM N.A.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     Chen A.C.;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               Pfam; PF00043;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         Transferase
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         Query Match
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    SEQUENCE
                                                           Query Match
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        Best Local
                                                                                                                                                                                                                                                                                                                                                   097117
                                                                                                                                                                                                                                                                                                                           RESULT 12
097117
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  QΩ
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            QY
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           δ
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            δ
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            δλ
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  g
                                                                                                                                                                                     g
                                                                                                                                                                                                                                     g
                                                                                                                                                                                                                                                                                                                                                                SOW
                                                                                                                                                             ŏ
                                                                                                                                                                                                                                                              δ
                                                                                                              δ
```

```
61 PYLIDGSHKVTQSNAILRYLGRKHNLCGETEEERIRVDILENRVMDTRIQMGMLCYXADF 120
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   58 PYYIDGDVKLTQSMAIIRYIADKHNMLGGCPKERAEISMLEGAVLDIRYGVSRIAYSKDF 117
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            118 ETLKVDFLSKLPEMLKMFEDRLCHKTYLNGDHVTHPDFMLYDALDVVLYMDPMCLDAFPK 177
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    INTESTINE. NOT EXPRESSED IN SPLEEN, KIDNEY, COLON, HEART, MUSCLE,
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           6; Gaps
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      01-JAN-1998 (TrEMBLrel. 05, Created)
01-JAN-1908 (TrEMBLrel. 13, Last sequence update)
01-JUN-2001 (TrEMBLrel. 17, Last annotation update)
01-JUN-2001 (TrEMBLrel. 17, Last annotation update)
01-JUN-2001 (TREMBLREL)
01-JUN-
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     De Bruin W.C.C., te Morsche R.H.M., Wagenmans M.J.M., Alferink J.C., Townsend A.J., Wieringa B., Peters W.H.M.; "Identification of a novel murine glutathione S-transferase class mu
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   4 LPI-LGYWKIKGLVQPTRLLLEYLEEKYEEHLYERDEG-----DKWRNKKFELGLEFPNL 57
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              -!- FUNCTION: CONJUGATION OF REDUCED GLUTATHIONE TO A WIDE NUMBER OF
                                                                                                                                                                                                                                                                                                                                                                                                                             Mus musculus (Mouse).
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         -!- SUBGNIT: HOMODIMER (BY SIMILARITY).
-!- SUBCELLULAR LOCATION: CYTOPLASMIC (BY SIMILARITY).
-!- TISSUE SPECIFICITY: EXPRESSED IN LIVER, STOMACH AND SMALL
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             11.0%; Score 479; DB 11; Length 218; 44.0%; Pred. No. 4.3e-23;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     EXOGENOUS AND ENDOGENOUS HYDROPHOBIC ELECTROPHILES (BY
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  -t- SIMILARITY: BELONGS TO THE GST SUPERFAMILY. MU FAMILY. BMDL; AJ000413; CAA04061.1; -. EMBL; AJ000412; CAA04060.1; -.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           Indels
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               PRINTS: PR01267; GSTRNSFRASEM.
Transferase; Multigene family.
SEQUENCE 218 AA; 25628 MW; 324B20CC6C35594B CRC64;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              -!- CATALYTIC ACTIVITY: RX + GLUTATHIONE = HX + R-S-G.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     SEQUENCE FROM N.A., AND TISSUE SPECIFICITY.
STRAIN=FVB, 129/SV, AND C57BL/6; TISSUE=LIVER;
MEDLINE=98145729; PubMed=9480867;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    178 LVCFKKRIEAIPQIDKYLKSSKYJAWPLQGWQATFG 213
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         218 AA.
218 AA.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           Mismatches
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          09TSM5;
01-MAY-2000 (TrEMBLrel. 13, Created)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    PRT;
           PRT;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                37;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 Biochem. J. 330:623-626(1998).
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       IPR003081; GST_mu.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           Conservative
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    PRELIMINARY;
           PRELIMINARY;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          InterPro; IPR000521; GST
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              MGD; MGI:1309467; Gstm6.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              GST; 1.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         Best Local Similarity
Matches 95; Conserv
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  P09488; 1GTU.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 BRAIN OR LUNG
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  NCBI_TaxID=10090;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 SIMILARITY)
                                                                                                                                                                                                                                                                                                                                                                                 GSTM6 OR GSTM5.
                                                            035660; 035661
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   Pfam; PF00043
                                                                                                                                                                                                                                                                                                                          CLASS M5).
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            InterPro;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             Query Match
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              Q9TSM5
```

2,

```
Wang C., Eaton D.L.;
"Mu-class GSTs Are Responsible for Aflatoxin B1-8,9-epoxide Activity
in the Non-human Primate Macaca fascicularis Liver.";
Submitted (NoV-1999) to the EMBL/GenBank/DDBJ databases.
EMBL: AP200709; AAF08539.1; -.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  62 DGDVKLTQSMAIIRYIADKHNMLGGCPKERAEISMLEGAVLDIRYGVSRIAYSKDFETLK 121
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      122 VDFLSKLPEMLKMFEDRLCHKTYLNGDHVTHPDFMLYDALDVVLYMDPMCLDAFPKLVCF 181
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   5; Gaps
                                                                                                                                                                                                                                                                                                                                                                                                                                                                      7 LGYWKIKGLVQPTRLLLEYLEEKYEEHLYERDEG-----DKWRNKKFELCLEFPNLPYYI 61
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        Macaca fascicularis (Crab eating macaque) (Cynomolgus monkey).
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Cercopithecidae;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  Macaca fuscata (Japanese macaque).

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Cercopithecidae;
Cercopitheinae; Macaca.
                                                                                                                                                                                                                                                                                                                                                                                                               Length 218;
01-MAY-2000 (TrEMBLrel. 13, Last sequence update)
01-JUN-2001 (TrEMBLrel. 17, Last annotation update)
GLUTATHIONE S-TRANSFERASE MU-CLASS SUBUNIT MI (EC 2.5.1.18).
                                                                                                                                                                                                                                                                                                                                                                                                                                            Indels
                                                                                                                                                                                                                                                                                                                                                             218 AA; 25577 MW; EE5B4D80F60C95EB CRC64;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   01-JUN-2001 (TrEMBLrel. 17, Created)
01-JUN-2001 (TrEMBLrel. 17, Last Sequence update)
01-JUN-2001 (TrEMBLrel. 17, Last annotation update)
GLUTATHIONE TRANSFERASE M2.
                                                                                                                                                                                                                                                                                                                                                                                                        11.0%; Score 476.5; DB 6;
43.9%; Pred. No. 6.2e-23;
tive 36; Mismatches 78;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            PRT; 221 AA
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               185 ISHFEGLEKISAYMKSSRFLPKPLYTRVAVWG 216
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            182 KKRIEAIPQIDKYLKSSKYIAWPLQGWQATFG 213
                                                                                                                                                                                                                                                                                                                               PRINTS; PR01267; GSTRNSFRASEM
                                                                                                                                                                                                                                                                                InterPro; IPR000521; GST.
InterPro; IPR003081; GST_mu.
Pfam; PF00043; GST; 1.
                                                                                                                                                                                                                                                                                                                                                                                                                          Best Local Similarity 43.9%
Matches 93; Conservative
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          PRELIMINARY;
                                                                                                              Cercopithecinae; Macaca
                                                                                                                                                        SEQUENCE FROM N.A.
                                                                                                                            NCBI_TaxID=9541;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 NCBI_TaxID=9542;
                                                                                                                                                                                                                                                                                                                                                Transferase.
                                                                                                                                                                                                                                                                                                                                                             SEQUENCE
                                                                                                                                                                                                                                                                                                                                                                                                            Query Match
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         Q9BEB0;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         Q9BEB0
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          RESULT 15
   QQ
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               Q
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             ò
```

3

6; Gaps

80; Indels

36; Mismatches

97; Conservative

Matches

Query Match Best Local Similarity

10.9%; Score 476; DB 6; Length 221; 44.3%; Pred. No. 6.8e-23;

Beuckmann C.T., Fujimori K., Urade Y.;
"Macaca fuscata glutathione transferase M2.";
Submitted (APR-1999) to the EMBL/GenBank/DDBJ databases.
EMBL; AB025799; BAB40442.1; -.
Transferase.
SEQUENCE 221 AA; 26068 WW; 01F22BAC6DA57FE9 CRC64;

SEQUENCE FROM N.A.

```
61 PNL/PYLIDGTHKITQSNAILRYIARKHNLCGETEKEKIREDILENQLMDNRMQLARLCYD 120
                                                                                                                                                                                                                        121 PDFEKLKPEYLEGLPEMLKLYSQFLGKQPWFLGDKITFVDF1AYDVLERNOVFEPSCLDA 180
                                                                                       55 PNLPYYIDGDVKLTQSMALIRYIADKHNMLGGCPKERAEISMLEGAVLDTRYGVSRIAYS 114
                                                                                                                                                                                    115 KDFETLKVDFLSKLPEMLKMFEDRLCHKTYLNGDHVTHPDFMLYDALDVVLYMDPMCLDA 174
                     1 MTKLPI-LGYWKIKGLVQPTRLLLEYLEEKYEEHLYERDEG-----DKWRNKKFELGLEF
                                                                                                                                                                                                                                                                                                         175 FPKLVCFKKRIEAIPQIDKYLKSSKYIAWPLQGWQATFG 213
                                         qq
                                                                                          ô
                                                                                                                                                                                    δŏ
                                                                                                                                                                                                                             qq
```

Search completed: January 15, 2002, 14:06:09 Job time: 948 sec

```
GenCore version 4.5 Copyright (c) 1993 - 2000 Compugen Ltd.
```

OM protein - protein search, using sw model

January 15, 2002, 14:06:31; Search time 134.53 Seconds (Without alignments) 319.662 Million cell updates/sec Run on:

1 OLGLVCLTVGPEVRFRTVTL......VEAKGKEEAIAALRLMAPFK 294 US-09-724-296-39 1527 Title: Perfect score: Sequence:

BLOSUM62 Gapop 10.0 , Gapext 0.5 Scoring table:

473505 seqs, 146272329 residues Searched:

473505 Total number of hits satisfying chosen parameters:

Minimum DB seq length: 0 Maximum DB seq length: 200000000

Post-processing: Minimum Match 0% Maximum Match 100% Listing first 45 summaries

sp_archea:*
sp_bacteria:*
sp_fungi:*
sp_human:*
sp_invertebrate:*
sp_mammal:* sp_organelle:*
sp_phage:*
sp_phage:*
sp_codent:*
sp_vorius:*
sp_virus:*
sp_virus:*
sp_virus:* SPTREMBL_17:* 10: 11: 12: 13: Database :

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

sp_unclassified:*

SUMMARIES

n	Description		O9s0m9 deinococcus	09rte6 deinococcus			-				Ogofia xvlella fas	O9vaq7 aeropyrim p	09cq24 lactococcus	091v03 arabidopsis	09x5n3 myxococcus	O9cab8 lactococcus	080962 arabidonsis	O9xac4 streptomyce	Ogrvo5 deinococcus	093161 mis miscoil	Q9kid6 streptomyce
SOUTHWALES	ID		CASCIMA	Q9RTE6	Q9L4F0	Q9L4E9	001408	Q9K9P8	010988	09AMN9	Q9PFJ8	09YAQ7	Q9CG24		Q9X5N3	Q9CGB8	080962	09XAC4	O9RYP5	_09JL61	09KID6
	DB		7	7	~	7	m	7	m	7	7	٦	7	10	7	7	10	7	7	11	7
	Query Match Length DB		300	326	317	325	929	322	599	315	437	313	1486	2216	526	535	791	368	515	658	388
æ	Query Match		y	99.2	19.5	19.5	17.2	16.4	15.1	7.3	6.7	6.5	6.5	6.3	6.3	6.3	6.3	6.2	6.2	6.3	6.2
	Score	1620	07CT	1520	297.5	297.5	262.5	250	230.5	111	102.5	98.5	98.5	96.5	96	96	96	95	95	95	94.5
	Result No.	-	-	2	m	4	S	9	7	80	6	10	11	12	13	14	15	16	17	18	19

			Q9adl6 polyangium Q915q9 pseudomonas Q91609 streptomyce P94134 alcaligenes Q9arl6 hordeum vul	Q9hxn5 pseudomonas Q9v696 drosophila Q918h4 streptomyce Q914w3 streptomyce Q918c9 polyanqium	054593 amycolatops Q9hz95 pseudomonas Q9kce3 bacillus ha Q9s2r3 streptomyce
2 09ACX5 2 085740 2 09HWG4 2 09Z3T9	2 09RI40 2 09RFM7 2 09RTX0 11 09ERV7	2 0942M0	2 Q9ADL6 2 Q915Q9 2 Q9FC09 2 P94134 10 Q9ARL6	2 Q9HXN5 5 Q9V696 2 Q9L8H4 2 Q9L4W3 2 Q9L8C9	2 054593 2 Q9H295 2 Q9KCE3 2 Q9S2R3
1203 1809 1809 2731	898 1809 603 915	932 777 951 1017		842 2040 4247 11096	3413 319 390 397
2222	00.1.1.0	0000		ພຸບບຸບ ລັລີລີລີ	დ. დ. დ. დ. დ. დ.
94.5 94.5 94.5	93.5 93.5 92.5	2222	91.5 91 90.5 90.5	800 800 800 800 800	90 89.5 89.5 89.5
222 232 332 332 332 332 332 332 332 332	24 25 27 27	30 30 31	2 2 2 2 2 2 2 2 2 2 2 2 2 2 2 2 2 2 2	37 38 39 40 41	42 44 45

ALIGNMENTS

RESULT 1 0950M9 AC 0950M9; DT 01-MAY-2000 (TYEMB DE UV-ENDONUCLEASE. GN UV-ENDONUCLEASE. GN UV-ENDONUCLEASE. GN UV-ENDONUCLEASE. GN UV-ENDONUCLEASE. GN UV-ENDONUCLEASE. AN ENDONUCLEASE. AN

ö

g

δ

```
16 RTVTLSRYRALSPAEREAKLLDLYSSNIKTLRGAADYCAAHDIRLYRLSSSLFPML---D 72
                                                                                                                                          SEQUENCE FROM N.A.
STRAIN=ATCC 14579 TYPE STRAIN;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                Query Match 19.5°
Best Local Similarity 27.1°
Matches 79; Conservative
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    PUTATIVE UV-ENDONUCLEASE.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             PRELIMINARY;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       NCBI_TaxID=1396;
                                                                                 NCBI_TaxID=1396;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    SEQUENCE
                                                                                                                                                                                                                                                                                                                                                                                                              SEQUENCE
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      Q9L4E9;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             Q9L4E9
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 RESULT
Q91.4E9
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         qq
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    Qy
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              ŏλ
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   δò
0
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     91 YRLSSSLFPMLDLAGDDTGAAVLTHLAPQLLEAGHAFTDAGVRLLMHPEQFIVLNSDRPE 150
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   121 VRESSVRAMSAHARVMDGLGLARTPWNLLLLHGGKGGRGAELAALIPDLPDPVRLRLGLE 180
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          181 NDERAYSPAELLPICEATGTPLVFDAHHHVVHDKLPDQEDPSVREWVLRARATWQPPEWQ 240
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  61 YRLSSSLFPMLDLAGDDTGAAVLTHLAPQLLEAGHAFTDAGVRLLMHPEQFIVLNSDRPE 120
190 NDERAYSPAELLPICEATGTPLVFDAHHHVVHDKLPDQEDPSVREWVLRARATWQPPEWQ 249
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             31 QLGLVCLTVGPEVRFRTVTLSRYRALSPAEREAKLLDLYSSNIKTLRGAADYCAAHDIRL 90
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           Gaps
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              1 QLGLVCLTVGPEVRFRTVTLSRYRALSPAEREAKLLDLYSSNIKTLRGAADYCAAHDIRL 60
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         White O., Elsen J.A., Heidelberg J.F., Hickey E.K., Peterson J.D., Dodson R.J., Haft D.H., Gwinn M.L., Nelson W.C., Richardson D.L., Moffat K.S., Qin H., Jiang L., Pamphile W., Crosby M., Shen M., Vamathevan J.J., Lam P., McDonald L., Utterback T., Zalewski C., Makarova K.S., Aravind L., Daly M.J., Minton K.W., Fleischmann R.D. Ketchum K.A., Nelson K.E., Salzberg S., Smith H.O., Venter J.C.,
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      Bacteria; Thermus/Deinococcus group; Deinococcales; Deinococcus.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           241 VVHLSNGIEGPQDRRHSHLIADFPSAYADVPQIEVEAKGKEEAIAALRLMAPFK 294
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          271 VVHLSNGIEGPQDRRHSHLIADFPSAYADVPWIEVEAKGKEEAIAALRLMAPFK 324
                                                                                    250 VVHLSNGIEGPQDRRHSHLIADFPSAYADVPWIEVEAKGKEEAIAALRLAMAPFK 303
                                                        241 VVHLSNGIEGPQDRRHSHLIADFPSAYADVPQIEVEAKGKEEAIAALRLMAPFK 294
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        0;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        "Genome sequence of the radioresistant bacterium Deinococcus
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          99.5%; Score 1520; DB 2; Length 326; 99.7%; Pred. No. 3.5e-114; Live 0; Mismatches 1; Indels
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        326 AA; 35693 MW; C4EA0D0AD2C38988 CRC64;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         01-OCT-2000 (TrEMBLrel. 15, Created)
U-OCT-2000 (TrEMBLrel. 15, Last sequence update)
01-MAR-2001 (TrEMBLrel. 16, Last annotation update)
PUTATIVE UV-ENDONUCLEASE.
                                                                                                                                                                                                                                                                                                                      01-MAY-2000 (TrEMBLrel. 13, Created)
01-MAY-2000 (TrEMBLrel. 13, Last sequence update)
01-MAR-2001 (TrEMBLrel. 16, Last annotation update)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    PRT; 317 AA
                                                                                                                                                                                                                                                                 326 AA
                                                                                                                                                                                                                                                                 PRT;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      MEDLINE=20036896; PubMed=10567266;
                                                                                                                                                                                                                                                                                                                                                                                                                 DAMAGE ENDONUCLEASE, PUTATIVE.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               Endonuclease; Complete proteome.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                EMBL; AE002022; AAF11370.1; -.
TIGR; DR1819; -.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       Science 286:1571-1577(1999).
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        Matches 293; Conservative
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    PRELIMINARY;
                                                                                                                                                                                                                                                                 PRELIMINARY;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                          Deinococcus radiodurans.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         Best Local Similarity
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              SEQUENCE FROM N.A.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  NCBI_TaxID=1299;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              radiodurans R1.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  Fraser C.M.;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           SEQUENCE
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                Query Match
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                Q9L4F0;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    Q9L4F0
                                                                                                                                                                                                                                                              Q9RTE6
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                YWJD.
                                                                                                                                                                                                        RESULT
                                                                                                                                                                                                                                 AND THE PROPERTY OF THE PROPER
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        RESULT
```

δλ g Ω Dp δ В δ a ò Q9L4F0

Op

ID AC DT DT DT DT DT DT DT DT

```
6
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   6
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                Okstad O., Gominet M., Purnelle B., Rose M., Lereclus D., Kolsto A.B.; "Sequence analysis of three Bacillus cereus loci under PICR-regulated genes encoding degradative enzymes and enterotoxin."; Microbiology 145:3129-3138(1999).
EMBL: AJ243712; CAB69814.1; '.
                                                                                                                                                                                                                                Okstad O., Gominet M., Purnelle B., Rose M., Lereclus D., Kolsto A.B.; "Sequence analysis of three Bacillus cereus loci under Pick-regulated genes encoding degradative enzymes and enterotoxin."; Microbiology 145:3129-3138(1999).
EMBL: AJ243712; CAB69813.1; -.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           190 ELLPICEATGTPLVFDAHHHVVH--DKLPDQEDPSVREWVLRARATWQ----PPEWQVVH 243
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       73 LAGDDTGAAVLTHLAPQLLEAGHAFTDAGVRLLMHPEQFIVLNSDRPEVRESSVRAMSAH 132
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                81 VEFDYIGV-----FTPLWRKIGALIKEHNLRISFHPNOFTLFTSDKPHITTNAITDMTYH 135
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      133 ARVMDGLGLARTPWNLLLLHGGKGGRGAELAAL --- IPDLPDPVRLRLGLENDERAYSPA 189
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        136 YKILDAIGIADSSYINIHVGGAYGNKEKAIERFHENIKKLPTHIKKQMTLENDDKTYTTA 195
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        Gaps
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              Gaps
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     16 RTVTLSRYRALSPAEREAKLLDLYSSNIKTLRGAADYCAAHDIRLYRLSSSLFPML---D 72
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  Query Match 19.5%; Score 297.5; DB 2; Length 325; Best Local Similarity 27.1%; Pred. No. 4.2e-16; Matches 79; Conservative 64; Mismatches 114; Indels 35;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       35;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         19.5%; Score 297.5; DB 2; Length 317; 27.1%; Pred. No. 4.1e-16;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   244 LSNGIEGPQDRRHSHLI-ADFPSAYADVPQ-----IEVEAKGKEEAIAAL 287
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               246 ISSPRSEKEFRAHAEYIDLEFIKPFLHIAKKHNHNFDIMIESKQKDLALFQL 297
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   64; Mismatches 114; Indels
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      325 AA; 37900 MW; 07EFBDD1FA580653 CRC64;
                                                                                                                                                                                                                                                                                                                                                                                                                                                          317 AA; 36990 MW; A6AC35F5800A22C3 CRC64;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          01-OCT-2000 (TrEMBLrel, 15, Last sequence update) 01-MAR-2001 (TrEMBLrel, 16, Last annotation update)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              Bacillus cereus.
Bacteria, Firmicutes, Bacillus/Clostridium group;
Bacteria, Firmicutes, Bacillus/Clostridium group, Bacillus/Staphylococcus group, Bacillus.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      325 AA
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    Bacillus/Staphylococcus group; Bacillus
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     01-OCT-2000 (TrEMBLrel. 15, Created)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                SEQUENCE FROM N.A.
STRAIN=ATCC 14579 TYPE STRAIN;
MEDLINE=20055637; PubMed=10589720;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      PRT;
                                                                                                                                                                                                              MEDLINE=20055637; PubMed=10589720;
```

```
13;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 "A eukaryotic gene encoding an endonuclease that specifically repairs DNA damaged by ultraviolet light."; EMBO J. 14:2393-2399(1995). EMBC j. 1392; CAB20113.1; -.
                                                                     89 VEFDYIGV----FTPLWRKIGALIKEHNLRISFHPNOFTLFTSDKPHITTNAITDMTYH 143
                                                                                                                         73 LAGDDTGAAVLTHLAPQLLEAGHAFTDAGVRLLMHPEQFIVLNSDRPEVRESSVRAMSAH 132
                                                                                                      ARVMDGLGLARTPWNLLLLHGGKGGRGAELAAL---IPDLPDPVRLRLGLENDERAYSPA 189
                                                                                                                                                                  190 ELLPICEATGTPLVFDAHHHVVH--DKLPDQEDPSVREWVLRARATWQ----PPEWQVVH 243
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       HDIRLYRLSSSLFPMLDLAGDDTGAAVLTHLAPQLL-EAGHAFTDAGVRLLMHPEQFIVL 114
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  292 YGIRFLRLSSEMFP---FASHPVHGYKLAPFASEVLAEAGRVAAELGHRLTTHPGQFTQL 348
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                115 NSDRPEVRESSVRAMSAHARVMDGLGLARTPWN-----LLLLHGGKGGRGAELAALI--- 166
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    167 ----POLPDPVRLRIGLENDERAYSPAELLPICEATGTPLVFDAHHHVV-----HDKL 215
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             216 PDQEDPSVREWVLRARATWQPPE-WQVVHLSNGIEG---PQDRR-HSHLIADFPSAYADV 270
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           39; Gaps
STRAIN=74-OR23-1A;
MEDLINE-95292980; PubMed=7774597;
Yajima H., Takao M., Yasuhira S., Zhao J.H., Ishii C., Inoue H.,
                                                                                                                                                                                                                                                                                                                                                                                                                                                            Eukaryota; Fungi; Ascomycota; Pezizomycotina; Sordariomycetes;
Sordariales; Sordariaceae; Neurospora.
                                                                                                                                                                                                                                             244 LSNGIEGPQDRRHSHLI-ADFPSAYADVPQ-----IEVEAKGKEEALAAL 287
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              Length 656;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           Indels
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              656 AA; 74413 MW; 4B231E7A8735C86F CRC64;
                                                                                                                                                                                                                                                                                                                                                                  Created)
Last sequence update)
Last annotation update)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          Query Match 17.2%; Score 262.5; DB 3; Best Local Similarity 32.2%; Pred. No. 6.9e-13; Matches 85; Conservative 38; Mismatches 102;
                                                                                                                                                                                                                                                                                                                                        656 AA
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      : :||| ||:|:
-DLMIEAKDKEQAV--FELMRTFK 541
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      271 PQIEVEAKGKEEAIAALRLMAPFK 294
                                                                                                                                                                                                                                                                                                                                        PRT;
                                                                                                                                                                                                                                                                                                                                                                  01-NOV-1996 (TrEMBLrel. 01, 01-MAX-1999 (TrEMBLrel. 10, 01-MAX-1999 (TrEMBLrel. 10,
                                                                                                                                                                                                                                                                                                                                       PRELIMINARY;
                                                                                                                                                                                                                                                                                                                                                                                                                                                 Neurospora crassa.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      SEQUENCE FROM N.A.
                                                                                                                                                                                                                                                                                                                                                                                                                 UV-ENDONUCLEASE.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         NCBI_TaxID=5141;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                Endonuclease
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              SEQUENCE
                                                                                                                                                                                                                                                                                                                                     001408
                                                                                                                                                                                                                                                                                                                                                     001408
                                                                                                                                                                                                                                                                                                        Ŋ
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  521
                                                                                                                                                                                                                                                                                                        RESULT
Q01408
              a
                                                                                                                                                                                                q
                                           ò
                                                                      ద
                                                                                                      ò
                                                                                                                                   q
                                                                                                                                                                  ŏ
                                                                                                                                                                                                                           ò
                                                                                                                                                                                                                                                          g
                                                                                                                                                                                                                                                                                                                                                    Q
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      g
```

RESULT

```
14;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        219 EDPSVREWVLRARATWQPPEWQV-VHLSNGIEGPQD-RRHSHLIADFPSAY----AD-V 270
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         105 LMHPEQFIVLNSDRPEVRESSVRAMSAHARVMDGLGLARTPWNLLLLH--GGKGGRGAEL 162
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               163 AALIPD----LPDPVRLRLGLENDERAYSPAELLPICEATGTPLVFDAHHH-VVHDKLPDQ 218
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      168 EQFIENTASIPKSLLSMIMLENDDKSYTIDDVLYLGEKLAIPVVLDIHHHDVLH----- 221
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             222 RSKSLQETWQRIVATWEDSPLPVKIHLSSPLSGEDDPRHHDYINADRFIAFLHEIGADAV 281
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  55 AHDIRLYRLSSSLFPML-----DLAGDDTGAAVLTHLAPQLLEAGHAFTDAGVRL 104
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    61 AYQISFFRLSSKLVPLVNHPLTEGWKYELA-----IAEELQAVGEFASEHQMRI 109
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              110 DFHPDHFVVLNSEAKEITRRSLQTLLYHYKLLKGMEI--DPRHRCVLHVGGKKKGVEAGL 167
                                                                                                                                                                                                                                                                                                                              "Complete genome sequence of the alkaliphilic bacterium Bacillus halodurans and genomic sequence comparison with Bacillus subtilis."; Mucleic Acids Res. 28:4317-4331(2000).
EMBL; AP001516; BAB06316.1: -.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         Gaps
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        1 QLGLVCLTV-----GPEVRFRTVTLSRYRALSPAEREAKLLD-LYSSNIKTLRGAADYCA 54
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           4 OFGYVAMSMELANASPS -- - KTMTATOFEK IEDHEAGLRKLERIAKTNLHNCLRLLKHNL 60
                                                                                                                                                                                                                                     STRAIN=C-125 / JCM 9153;
MEDLINE=20512582: PubWed=11058132;
Takami H., Nakasone K., Takaki Y., Maeno G., Sasaki R., Masul N.,
Fuji F., Hirama C., Nakamura Y., Ogasawara N., Kuhara S.,
Horikoshi K.;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         52;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                               16.4%; Score 250; DB 2; Length 322; 26.2%; Pred. No. 2.7e-12; Live 68; Mismatches 116; Indels
                                                                                                                                                                                                                                                                                                                                                                                                                         322 AA; 36900 MW; 4B0DC9291CD1E7B4 CRC64;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              Eukaryota, Fungi, Ascomycota, Schizosaccharomycetes, Schizosaccharomycetales, Schizosaccharomycetaceae;
                              01-0CT-2000 (TrEMBLrel. 15, Created)
01-0CT-2000 (TrEMBLrel. 15, Last sequence update)
01-0CT-2000 (TrEMBLrel. 15, Last annotation update)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   Last sequence update)
Last annotation update)
                                                                                                                             Bacillus halodurans.
Bacteria; Firmicutes; Bacillus/Clostridium group;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 599 AA
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              Schizosaccharomyces pombe (Fission yeast).
                                                                                                                                                                Bacillus/Staphylococcus group; Bacillus
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 Q10V80
Q10988; P87339;
Q1-NOV-1996 (TERMBLrel. 01, Created)
Q1-NOV-1996 (TERMBLrel. 01, Last sequal-NOV-1998 (TERMBLrel. 08, Last ann
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               PRT;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 271 PQIEVEAKGKEEAIAALRLM 290
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       282 DHLHVMIEAKKKDLALFQLM 301
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  Best Local Similarity 26.29
Matches 84; Conservative
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               PRELIMINARY;
PRELIMINARY;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   Schizosaccharomyces.
                                                                                                                                                                                                                      SEQUENCE FROM N.A.
                                                                                                                                                                                                                                                                                                                                                                                                        Complete proteome. SEQUENCE 322 AA;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           SEQUENCE FROM N.A.
                                                                                                                                                                                   NCB1_TaxID=86665;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        UV-ENDONUCLEASE.
UVDE OR UVE1+.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        NCB1_TaxID=4896;
                                                                                          PROTEIN.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            STRAIN=SP972
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   Query Match
                 Q9K9PB;
                                                                                        BH2597
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               010988
Q9K9P8
                                                                                                              BH2597
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            010988
q
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         ŏλ
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   q
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             q
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 δλ
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                AC OC OC OX OX RAP REP REP
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 ΟY
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          ò
```

```
10;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              249 RLGYACLNTILRSMKERVFCSRTCRITTIQRDGLESVKQLGTQNVLDLIKLVEWNHNFGI 308
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     59 RLYRLSSSLFPMLDLAGDDTGAAVLTHLAPQLLEAGHAFTDAGVRLLMHPEQFIVLNSDR 118
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             309 HFMRVSSDLFP---FASHAKYGYTLEFAQSHLEEVGKLANKYNHRLTMHPGQYTQIASPR 365
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              119 PEVRESSVRAMSAHARVMDGLGLARTPWN-----LLLLHGGKGGRGAE----LAALI 166
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                36 EVVVDSAIRDLAYHDEI----LSRMKLNEQLNKDAVLIIHLGGTFEGKKETLDRFRKNY 420
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       167 PDLPDPVRLRLGLENDERAYSPAELLPICEATGTPLVFDAHHHVV------HDKLPD 217
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           421 QRLSSDSVKARLVLENDDVSWSVQDLLPLCQELNIPLVLDWHHHNIVPGTLREGSLDLMP- 479
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 218 QEDPSVREWVLRARATWQPPEWQVVHLS----NGIEGPQDRRHSHLIADFPSAYADVPQ 272
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          480 -LIPTIRETWTRKGIT-----QKQHYSESADPTAISGMKRRAHSDRVFDFPPCDPTM-D 531
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     Gaps
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          1 QLGLVCLTVGPEVRFRTVTLSRYRALSPAEREA--KLLDLYSSNIKTLRGAADYCAAHDI 58
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           Tan B.-H., Yeo C.-C., Poh C.-L.;
"Pseudomonas alcaligenes NCIB 9867 putative heat shock protein and truncated group II intron.";
Submitted (NOV-2000) to the EMBL/GenBank/DDBJ databases.
EMBL: AF323438: AAK11471.1;
                Takao M., Yonemasu R., Yamamoto K., Yasui A.;
Characterization of a UV endonuclease gene from the fission yeast
Schizosaccharomyces pombe and its bacterial homolog.";
Nucleic Acids Res. 24:1267-1271(1996).
                                                                                                                                                                                                                                                                                     Lyne M., Rajandream M.A., Barrell B.G., Lucas M., Gaillardin C., Submitted (JUN-1998) to the EMBL/GenBank/DDBJ databases.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     42; Mismatches 140; Indels 45;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                Bacteria; Proteobacteria; gamma subdivision; Pseudomonadaceae;
                                                                                                                                                                                                                                                                                                                                                                                                                                                               Score 230.5; DB 3; Length 599; Pred. No. 2.3e-10;
                                                                                                             SEQUENCE FROM N.A.
MEDLINE=97175806; PubMed=9023111;
Davey S., Nass M.L., Ferrer J.V., Sidik K., Eisenberger A., Mitchell D.L., Freyer G.A.,
                                                                                                                                                                                        "The fission yeast UVDR DNA repair pathway is inducible.", Nucleic Acids Res. 25:1002\cdot1008(1997).
                                                                                                                                                                                                                                                                                                                                                                                                        599 AA; 68816 MW; C09ED7219D5895BE CRC64;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   01-JUN-2001 (TrEMBLrel. 17, Last sequence update) 01-JUN-2001 (TrEMBLrel. 17, Last annotation update)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                315 AA
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  01-JUN-2001 (TrEMBLrel. 17, Created)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                PRT;
MEDLINE=96188860; PubMed=8614629;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         HYPOTHETICAL 34.9 KDA PROTEIN.
                                                                                                                                                                                                                                                                                                                                                                    EMBL; AL023859; CAA19577.1; -.
                                                                                                                                                                                                                                                                                                                           EMBL; D78571; BAA11415.1; -.
EMBL; U78487; AAC49664.1; -.
                                                                                                                                                                                                                                                                                                                                                                                                                                                               15.1%;
27.9%;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     88; Conservative
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             532 LMIEAKEKEQAVFEL 546
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          273 IEVEAKGKEEAIAAL 287
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                PRELIMINARY;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               Pseudomonas alcaligenes
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   Local Similarity
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        SEQUENCE FROM N.A.
                                                                                                                                                                                                                                                  SEQUENCE FROM N.A.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    NCBI_TaxID=43263;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               STRAIN=NCIB 9867;
                                                                                                                                                                                                                                                                       STRAIN=972H-;
                                                                                                                                                                                                                                                                                                                                                                                          Endonuclease
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  Pseudomonas
                                                                                                                                                                                                                                                                                                                                                                                                          SEQUENCE
                                                                                                                                                                                                                                                                                                                                                                                                                                                                   Query Match
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  09AMN9;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              99AMN9
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        æ
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       Matches
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     RESULT
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              O9AMN9
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     ολ
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       d
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              δy
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                pp
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       δy
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         qq
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 δ
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    Dp
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          ŏ
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 g
```

```
13;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               RECONTING—20365717; PubMed=10910347;

RA SIMPSON A.JG., Reinach F.C., Arruda P., Abreu F.A., Acencio M.,

RA Alvarenga R., Alves L.M.C., Arraya J.E., Bala G.S., Baptista C.S.,

RA BALVARENGA R.H. C., Arraya J.E., Bove J.M., Briones M.R.S.,

RA BALVA COROMDO C., Costa M.C. Carraro D.M., Carrer H.,

RA Golauto N.B., Colombo C., Costa M.C.R., Costa NCC. El-Dorry H.,

RA Fadincani A.P., Ferreira A.J.S., Ferreira V.C.A., Ferro J.A.,

RA Fadincani A.P., Ferreira A.J.S., Ferreira V.C.A., Ferro J.A.,

RA Garnier M., Goldman G.H., Goldman M.H.S., Gomes S.L., Gruber A.,

RA Garnier M., Goldman G.H., Goldman M.H.S., Gomes S.L., Kitajima J.P.,

RA Garnier M., Goldman G.H., Goldman M.H.S., Momes E.E., Kitajima J.P.,

RA Garnier M., Madeira A.M.B.N., Madeira H.M.F., Marino C.L.,

RA Lemos E.G.M., Lemos M.V.F., Lopes C.R., Machado J.A.,

RA Machado M.A., Madeira A.M.B.N., Madeira H.M.F., Marino C.L.,

RA Machado M.A., Madeira A.M.B.N., Mattins E.M.F., Matsukuma A.Y.,

RA Machado M.A., Mattins E.A.L., Martins E.M.F., Matsukuma A.Y.,

RA Moon D.H., Nasaimento A.L.T.O., Netto L.E.S.,

RA Moon D.H., Nasaimento A.L.T.O., Netto L.E.S.,

RA Goliveira M.C., de Oliveira R.C., Pelmileri D.A., Paris A.,

RA de Silveira J.F., Silvestri M.L.Z., Siqueira W.J., de Souza A.A.,

RA de Souza A.P., Terenzi M.F., Traai S.N., Tsuhako M.H.,

RA de Sulva A.C.R., da Silva A.M., da Silva F.R., Silva W.A. J.,

RA de Souza A.P., Terenzi M.F., Trais S.N., Tsuhako M.H.,

RA Zago M.A., Zatz M., Meidanis J., Setubal J.C.;

RA The genome sequence of the plant pathogen Xylella fastidiosa.";
                                                                                                                                                                                                                                                                                                                                                                                                                    101 KTVNKSRGGAVLLKEP---ELLH----AIVREVRRTVPAEIPVTAKMRLGFEGKEGALDC 153
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              48 KLAPELFNGSR--TRAGT----PMRVQFLGSDPQCLADNAAFACTLGAPVIDLNFGCPA 100
                                                                                                                                                                                                                                                                                                                                                                                   138 -----GLGLARTPWNLLLLHGGKGGRGAELAALIP-DLPDPVRLRLGLENDERAYSP 188
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     189 AELLPICEATGTPLVFDAHHHVVH--DKLPDQEDPSVREWVLRAR------ATWQ 235
                                                                                                                                                                                                                                                                                         85 HLAPQILLEAGHAFTDAGVRLLMHPEQFIVLNSDRPEVRESSVRAMSAHARVMD----- 137
                                                                                                                                                                                                                                     4 ALAPME---GLVDEILRDVLTRIGGIDWCVTEFI---RVSERLLP------AATYH 47
                                                                                                                                                                                             25 ALSPAEREAKLLDLYSSNIKTLRGAADYCAAHDIRLYRLSSSLFPMLDLAGDDTGAAVLT 84
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       Bacteria; Proteobacteria; gamma subdivision; Xanthomonas group;
                                                                                                                                                68;
                                                                                            Query Match 7.3%; Score 111; DB 2; Length 315; Best Local Similarity 25.2%; Pred. No. 0.37; Matches 64; Conservative 30; Mismatches 92; Indels 64
Hypothetical protein.
SEOUENCE 315 AA; 34922 MW; 07AF8DD29EEAB90D CRC64;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  Last sequence update)
Last annotation update)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              437 AA
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              Created)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            PRT;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         01-OCT-2000 (TrEMBLrel. 15, 01-OCT-2000 (TrEMBLrel. 15, 01-MAR-2001 (TrEMBLrel. 16,
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 PRELIMINARY;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              236 PPEWQVVHLSNGIE 249
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              205 LDDWRRCREISGVD 218
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                Xylella fastidiosa.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              CYCLE PROTEIN.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 SEQUENCE FROM N.A.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 NCBI_TaxID=2371;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          STRAIN=9A5C
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             Xylella
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       Q9PFJ8;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              Q9PFJ8
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       XF0659
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    9
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    RESULT
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          Q9PFJ8
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              δŽ
                                                                                                                                                                                                                                                                                                                                                                                                                                     q
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     δ
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               qq
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              Qγ
```

RL DR KW SQ	Nature 406:151-159(2000). EMBL; AE003910; AAF83469.1; Complete proteome. SEQUENCE 437 AA; 48811 MW; ABFBE0825B819DF2 CRC64;	Qy	159GAELAALIF : : 210 EIEEARGSSSVEAYST
Ōmű	<pre>Query Match 6.7%; Score 102.5; DB 2; Length 437; Best Local Similarity 25.1%; Pred. No. 2.8; Matches 56; Conservative 30; Mismatches 84; Indels 53; Gaps 12;</pre>	Qy	212 HDKLPDQEDPSVREWV : : : 263MASVEEPSYRD-V
Qy Db	43 IKTLRGAADYCAAHDIRLYRLSSSLFPMLDLAGDDTGAAVLTHLAPQLLEAGHAF 97 ::	RES 090 ID	RESULT 11 09CG24 ID 09CG24 PRELIMINA
Oy Dp	98 TDAGVRLLMHPEQFIVLNSDRPEVRESSVRAMSAHARVMDGLGLARTPWNLLLLHGGKGG 157 17 EDSSNADLRHDRNFLRIHV-LPLLHORWPQATAVLARNAALAAANADLL 224	AC DT DT DT	09CG24; 01-JUN-2001 (TrEMBLre 01-JUN-2001 (TrEMBLre 01-JUN-2001 (TrEMBLre
QY DP	158 RGAELAALIPDLPDPVRLRLGLE-NDERAYSPAELLPICEATGTPLVFDAHHHVV- 211	G C C C C C C C C C C C C C C C C C C C	GLUTAMATE SYNTHASE LA GLTB. Lactococcus lactis (9 Bacteria; Firmicutes;
Qy Dp	212 HDKLPDQEDPSVREWVLRARATWQPPEWQ 240 :	OC OX RR RP RP	Lactococcus. NCBL_TaxID=1360; [1] SEQUENCE FROM N.A.
RESULT Q9YAQ7	10 NAAAA DEELTMINAEV. DEET.	RA RA RT	Bolotin A., Wincker F Weissenbach J., Ehrli "The complete genome
1 2 E	;; -1999 (TrEMBLrel. 12, -1609 (PrEMBLrel 12,	R R T	<pre>lactis."; Genome Res. 0:0-0(200 EMBL; AE006360; AAK05 Interpret</pre>
OOG	OI NOV 1999 (ILEMBLET. 14, Last Sequence update) 01-MAY-2001 (TrEMBLTel. 16, Last annotation update) HYPOTHETICAL 34.0 KDA PROTEIN APE1886. APE1886.	2	InterPro; IPR002489; InterPro; IPR002932; Pfam; PF01493; DUF14; Dfam: DF01645; Clu sv
8888	Acropyrum pernix. Archaea; Crenarchaeota; Desulfurococcales; Desulfurococcaceae; Aeropyrum.	KW SQ	Complete proteome. SEQUENCE 1486 AA;
N R R R	NCBI_TAXID=50550; [1] SEQUENCE FROM N.A. STRAIN=K1:	υ ш 2	Query Match Best Local Similarity Matches 66. Conservat
8 % §	MEDLINE=99310339; Pubmed=10382966; Kawarabayasi Y., Hino Y., Horikawa H., Yamazaki S., Haikawa Y.,	oy ,	RFRTVTI
RA RA	Jin-no K., Takahashi M., Sekine M., Baba SI., Ankai A., Kosugi H., Hosoyama A., Fukui S., Nagai Y., Nishijima K., Nakazawa H., Tabamiya M. Masuda G. Dunahashi M. Tabana M. Wadak	qq	: 233 RFSTNTFPSWNRAQPE
RA R	Yamazaki J., Kushida N., Oguchi A., Aoki KI., Kubota K., Nakamura Y., Nomura N., Sako Y., Kikuchi H.;	Oy	62RLSSSI
RT RL	"Complete genome sequence of an aerobic hyper-thermophilic crenarchaeon, Aeropyrum pernix Kl."; DNA Res. 6:83-101(1999).	DP Qy	288 CMEYFYRNGRELPESI 100 -AGVRL
S K	EMBL; AP000062; BAA80891.1; Hypothetical protein; Complete proteome. SEQUENCE 313 AA; 33960 MW; A796826413607CD0 CRC64;	qq	 348 TVGARLDRNGLRPSR\
Č		Qy	142 ARTPWNLLLLHGGKGC
Z a c	Query march 8.3%; Score 98.3; DB 1; Length 313; Best Local Similarity 25.1%; Pred. No. 3.7; Matches 65; Conservative 35; Mismatches 84; Indels 75; Gaps 15;	DB Qy	395PGNMLLVNTDEG
Qy	11 PEVRERTUTLSRYRALSP-ABEREAKLLDLYSSNIKTLRGAADDYCAAHDIRLYRLSSS 66	qa	
Qy Db	67 LFPMLDLAGDDTGAAVLTHLAPQLLEAGHAFTDAGVRLLMHPEQFTVLNSDRPE 120 106 EFTLVAFTGDPKDSRIV-HLAEAASLLGARGLYVVGPPMHPGYEERLSMLGASRVE 160	A I D	SULF 12 2V03 Q9LV03; Q9LV03;
Qy Db	121 V-RESSVRAMSAHARVMDGLGLARTPWNLLLLHGGKGGR	DI	01-0CT-2000 (Tremble 01-0CT-2000 (Tremble 01-JUN-2001 (Tremble NADH-DEPENDENT GLUTAN

```
P., Mauger S., Jaillon O., Malarme K., ich S.D., Sorokin A.; sequence of the lactic acid bacterium Lactococcus
                    6.5%; Score 98.5; DB 2; Length 1486;
23.2%; Pred. No. 29;
Live 33; Mismatches 81; Indels 105; Gaps
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       3--RGAELAALIPD-LPDPVRLRLGL-----ENDERAYSPAE- 190
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  KIIRNEEVKSYYANKYPYQEFLSAGLKKLSALTESEKTNPIPSAK 451
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     LMHPEQFIVLNSD------RPEVRESSVRAMSAHARVMDGLGL 141
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           DLPDPVRLRLGLENDERAYSPAELLPICEATGTPLVFDAHHHVV 211
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        JEPMLDLA-GDDTGAA------91THLAPQLLEAGHAFTD-- 99
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               AEREAKLIDLYSSNIKTLRGAADYCAAHDIRLY------ 61
                                                                                                                                                                                                                                                                                                                                                  subsp. lactis) (Streptococcus lactis).
; Bacillus/Clostridium group; Streptococcaceae;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       164390 MW; 0E9C715CD3C4940D CRC64;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           ---LLPICEATGTPLV---FDAHHHVVHDK 214
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     rel. 15, Created)
rel. 15, Last sequence update)
rel. 17, Last annotation update)
WATE SYNTHASE.
                                                                                                                                                                                                                                                   el. 17, Created)
el. 17, Last sequence update)
el. 17, Last annotation update)
ARGE SUBUNIT (EC 1.4.1.13).
                                                                                                                                                                                                                PRT; 1486 AA.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    PRT; 2216 AA.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            Glu_synthase.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   01).
5384.1; -.
DUF14.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    nthase; 1.
                                                                                                 |:|:
VIRS 277
                                                                        'LRA 230
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       ARY;
```

14;

OLIGOPEPTIDE PERMEASE HOMOLOG OPPA.

```
δλ
                                                                                                                                                                                                                                                                                                                                                                   qq
                                                                                                                                                                                                                                                                                                                                                                                              ōλ
                                                                                                                                                                                                                                                                                                                                                                                                                      qq
                                                                                                                                                                                                                                                                                                                                                                                                                                                 Óλ
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        17;
                                                                                                                                              "Structural analysis of Arabidopsis thaliana chromosome 5. X. Sequence features of the regions of 3,076,755 bp covered by sixty Pl and TAC
         Eukaryota; Viridiplantae; Streptophyta; Embryophyta: Tracheophyta;
Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; Rosidae;
eurosids II; Brassicales; Brassicaceae; Arabidopsis.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          332 RFSTNTFPSWDRAQP----MRVLG-HNGEINTLRGNVNWMRAREGLLKCNELGLSKKELK 386
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        67 -LFPMLDLAGDDTGAAVLTHLAPQLLEAGHAFTDAGVRLLMHPEQFIVLNSDR--PEVRE 123
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             124 SSVRAMSAHARVMDGLGL-------ARTPWNLLLLHGGKGGRGAELAALIP 167
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  168 DLPDPVRLRLGLENDERAYSPAELLPICEATGTPLVFDAHHHVVHDKLPDQEDPSVREWV 227
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            498 DVPPEDVMRKGRLN-------PGMMLLVDFEKHIVVD------DDALKQQY 535
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     228 LRARATWQPPEWQVVHLSNGIEG-PQDRRHSHLIADFPSAYADVPQ---IEVEAKGKEEA 283
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        440 EFYEYLSALMEPWDGPALISFTDGRYLGATLDRNGLRPGRFYITHSGRVIMASEVG--VV 497
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              Gaps
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     14 RFRTVTLSRYRALSPAEREAKLLDLYSSNIKTLRGAADYCAAHD--IRLYRLSSS----- 65
                                                                                                                                                                                      DNA Res. 7:31-63(2000).
-1- COFACTOR: FAD (BY SIMILARITY).
-1- SIMILARITY: TO PYRIDINE NUCLEOTIDE-DISULPHIDE OXIDOREDUCTASES
                                                                                                   MEDLINE-20181125; PubMed-10718197;
Sato S., Nakamura Y., Kaneko T., Katoh T., Asamizu E., Kotani H.,
Tabata S.;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             75;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                Query Match
6.3%; Score 96.5; DB 10; Length 2216;
Best Local Similarity 22.0%; Pred. No. 72;
Matches 69; Conservative 54; Mismatches 116; Indels 75;
                                                                                                                                                                                                                                                                                                                                                                          Pfam; PF00070; pyr_redox; 1.

PRINTS; PR00419; ADXRDTASE.

PRINTS; PR00119; FADPRI.

PRINTS; PR00469; PNDRDTASEI.

PRINTS; PR00469; PNDRDTASEII.

FAD; Flavoprotein; Oxidoreductase; Redox-active center.

SEQUENCE 2216 AA; 242877 MW; 2AF258AB8C682371 CRC64;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              01-NOV-1999 (TrEMBLrel. 12, Created)
01-NOV-1999 (TrEMBLrel. 12, Last sequence update)
01-JUN-2001 (TrEMBLrel. 17, Last annotation update)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       526 AA.
Arabidopsis thaliana (Mouse-ear cress).
                                                                                                                                                                                                                                                                               Interpro; IPR001327; FAD_pyr_redox.
InterPro; IPR002932; Glu_synthase.
InterPro; IPR000205; NAD_binding.
InterPro; IPR001100; pyr_redox_2.
                                                                                                                                                                                                                                        EMBL, AB022754; BAA97323.1; -.
InterPro; IPR000759; Adrndx_reductase.
InterPro; IPR002489; DUF14.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          PRT;
                                                                                                                                                                                                                                                                                                                                               Pfam; PF01493; DUF14; 1.
Pfam; PF01645; Glu_synthase; 1.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     PRELIMINARY;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          284 IAAL---RLMAPFK 294
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    590 MESMGIHGLLSPLK 603
                                                                             SEQUENCE FROM N.A.
                                                   NCBI_TaxID=3702;
                                                                                                                                                                                                                              CLASS-I
                                                                                                                                                                            clones.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    Q9X5N3;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     09X5N3
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             RESULT 13
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         09X5N3
 ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 Op
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             δy
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       Op
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         Dp
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            qq
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          óγ
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        ID
DI
DI
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               QΥ
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     ŏ
```

```
17;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         Bolotin A., Wincker P., Mauger S., Jaillon O., Malarme K., Weissenbach J., Ehrlich S.D., Sorokin A.;
"The complete genome sequence of the lactic acid bacterium Lactococcus
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          105 LMH-PEQ-----FIVLNS-----DRPEVRESSVRAMSAHARVMDGLGLARTPWNL 148
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 135 VTLARPRSYFLA-RVANVYLFYPAPSADLEGKSDEEVRDYFDRPRDGRPLALGPYRVERW 193
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       53 -CAAHDIRLYRLSSSLFPMLDLAGDDTGAAVLT----HLAPQLLE---AGHAFTDAGVRL 104
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    88; Gaps
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     Gaps
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          11 PEVRFRTVTLSRYRALSPAEREAKLLD--LYSSN1KTL-RGAADYCAAHDIRL---YRLS 64
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          Lactococous lactis (subsp. lactis) (Streptococcus lactis).
Bacteria; Firmicutes; Bacillus/Clostridium group; Streptococcaceae;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  54;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    Complete protecme; FAD; Flavoprotein; Hypothetical protein;
Myxococcus xanthus.
Bacteria; Proteobacteria; delta subdivision; Myxobacteria;
Myxococcales; Cystobacterineae; Myxococcaceae; Myxococcus.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              Query Match 6.3%; Score 96; DB 2; Length 535; Best Local Similarity 24.4%; Pred. No. 12;
                                                                                                                                                                                                                                                                                                                                                                                                                            6.3%; Score 96; DB 2; Length 526;
26.6%; Pred. No. 12;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    36; Mismatches 87; Indels
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     81; Indels
                                                                                                                                                                                            Ueki T., Inouye S.;
Submitted (FEB-1999) to the EMBL/GenBank/DDBJ databases.
                                                                                                                                                                                                                                                  EMBL, AF127082; AAD31006.1; -.
Interpro; IPR000914; SBP_bac_5.
Pram; PF00496; SBP_bac_5; 3.
SEQUENCE 526 AA; 58593 MW; AFAS85E37B207DB3 CRC64;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          535 AA; 58670 MW; 116102ABB9FAEC95 CRC64;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             01-JUN-2001 (TrEMBLrel. 17, Created)
01-JUN-2001 (TrEMBLrel. 17, Last sequence update)
01-JUN-2001 (TrEMBLrel. 17, Last annotation update)
HYPOTHETICAL PROTEIN YLJF.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        18 VTLSRYRALSPAEREAKLLDLYSSNIKTLRGAADY-----
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    312 LLPPELPGAATPEQAARLPRYAPEQARAELAGVERPLR 349
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              149 LLLHGGKGGRGAELAALIPD-LPDPVRLRL-GLENDER 184
                                                                                                                                                                                                                                                                                                                                                                                                                                           26.6%; Pred. No. +ive 25; Mismatches
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              EMBL, AE006350, AAK05277.1; -. InterPro; IPR001327; FAD_PYr_redox. PRINTS; PR00368; FADPNR.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                PRT;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     Genome Res. 0:0-0(2001).
-!- COFACTOR: FAD (BY SIMILARITY).
                                                                                                                                                                                                                                                                                                                                                                                                                                                       Best Local Similarity 26.6
Matches 58; Conservative
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       Conservative
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          PRELIMINARY;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      SEQUENCE FROM N.A.
                                                                                                                                      SEQUENCE FROM N.A.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              NCBI_TaxID=1360;
                                                                             NCBI_TaxID=34;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  Oxidoreductase.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    STRAIN-IL1403;
                                                                                                                                                                       STRAIN=DZF1;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   lactococcus.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       68;
                                                                                                                                                                                                                                                                                                                                                                                                                               Query Match
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                lactis.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          09CGB8
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        RESULT 14
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       Matches
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  Q9CGB8
```

```
18;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          281 QVYRGRTLDGADVAVKVQRPDLRHAVLRDIYILRLGLGVLRKVAKRENDIRVYADELGMG 340
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       LEVFSAFTIAAIRLRTSASVKGKSLEASEQIFGMVLKETMLHLGPTFIKVGQSLSTRPDI 226
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               158 ------RGAELAALI--PDLPDPVR-----LRLGL------ENDERAYS----- 187
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           188 -PAELLPICEATGTPLVFDAHHVVHDKLPDQEDPSVREWVLRARATWQPPEWQVVHLSN 246
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                341 LAGELDFTLEAANASEFQEAHARFSYIRVPKVYQHLTRKRVLTM------EWMVG---- 389
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          LEAGHAFTDAGVRL-----SDRPEQ------FIVLN---SDRPEV 121
                                                                          69 SKKIKNVSLAPELDYVNPVIGSKKMLH-RPVVI----GFGPAG----MFAALLLAQNGYR 119
                                                      65 S-----SLFPMLDLAGDDTGAAVLTHLAPQLLEAGHAFTDAGVRLLMHPEQFIVLNSDR 118
                                                                                                                                 119 PEVRESSVRAMSAHARVMDGLGLARTPWNLLLLHG------GKGGRG----AELAA 164
                                                                                                                                                         165 LIPDLPDPVRLRLGLENDERAYSPAELLPICEATGTPLVFDAHHHVVHDKLPD----- 217
                                                                                                                                                                                                                                         Arabidopsis thaliana (Mouse-ear cress).
Eukaryota, Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; Rosidae;
eurosids II; Brassicales; Brassicacee; Arabidopsis.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  Ouery Match 6.3%; Score 96; DB 10; Length 791;
Best Local Similarity 24.1%; Pred. No. 20;
Matches 83; Conservative 30; Mismatches 107; Indels 124; Gaps
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             36 LDLYSSNIKTLRGAADYCAAHDIRLYRLSSSLFPMLDLAGDDTGAAVLTHLAP----QL 90
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             SEQUENCE FROM N.A.
STRAIN=CV. COLUMBIA;
Rounsley S.D., Kaul S., Lin X., Ketchum K.A., Crosby M.L.,
Brandon R.C., Sykes S.M., Mason T.M., Kerlavage A.R., Adams M.D.,
Somerville C.R., Venter J.C.;
"Arabidopsis thaliana chromosome II BAC T16B24 genomic sequence.";
Submitted (AUG-1998) to the EMBL/GenBank/DDBJ databases.
EMBL, AC004697: AAC28988 1; -
EMBL, AC004697: PROFICIASE_ASP; UNKNOWN 1.
SEQUENCE 791 AA; 88694 MW; C771BC05757D2863 CRC64;
14 PVSKVKSLVLKKLRL----QESDLLDYRIXKESIDARHRGEIDFIYTVDIKLKDEARIL
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      247 GIEGPODRRHSHLIADFPSAYADVPQIEVEAKGKEEAIAALRLM 290
                                                                                                                                                                                                                                                                                                                                                                                                                                                               01-NOV-1998 (TrEMBLrel. 08, Created)
01-NOV-1998 (TrEMBLrel. 08, Last sequence update)
01-JUN-2001 (TrEMBLrel. 17, Last annotation update)
T16B24.17 PROTEIN.
                                                                                                                                                                                                                                                                                                                  214 QIIELGGEVHFDAQVEEFLI-----EDEELQAVKLADG 246
                                                                                                                                                                                                                                                                                        -----QEDPSVREWVLRARATWQPPEWQVVHLSNG 247
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    122 -- RESSVRAMSAHARVMDGLGLARTPW--NLLLLHGGKGG-
                                                                                                                                                                                                                                                                                                                                                                                                                           791 AA
                                                                                                                                                                                                                                                                                                                                                                                                                             PRT;
                                                                                                                                                                                                                                                                                                                                                                                                                             PRELIMINARY;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            NCBI_TaxID=3702;
                                                                                                                                                                                                                                                                                                                                                                                                                             080962
                                                                                                                                                                                                                                                                                                                                                                                    RESULT 15
080962
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          91
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              167
                                                                                                                                                                                                                                                                                          218
                                                                                                                                                                                                                                                                                                                                                                                                                             ŏ
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   q
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          õ
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          q
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 ŏ
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         ŏ
                                                                                            g
                                                                                                                                                                                                                                               g
                                                                                                                                                                                                                                                                                          ò
                                                        ò
                                                                                                                                   õ
                                                                                                                                                                     g
                                                                                                                                                                                                           οy
```

```
390 --ESPTD-----LLA-ISSGYSD---HDNESHEKQKIEARRRLL 422
                       qq
```

Search completed: January 15, 2002, 14:06:33 Job time: 972 sec

.

GenCore version 4.5 Copyright (c) 1993 - 2000 Compugen Ltd.

using sw model OM protein - protein search, Run on:

January 15, 2002, 13:52:37; Search time 74.94 Seconds (without alignments) 298.843 Million cell updates/sec

US-09-724-296-39 1527 1 QLGLVCLTVGPEVRFRTVTL.....VEAKGKEEAIAALRLMAPFK Perfect score: Sequence:

219241 seqs, 76174552 residues BLOSUM62 Gapop 10.0 , Gapext 0.5 Searched:

Scoring table:

219241 Total number of hits satisfying chosen parameters:

Minimum DB seq length: 0 Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0% Maximum Match 100% Listing first 45 summaries

PIR_68:* Database

1: pirl:* 2: pir2:* 3: pir3:* 4: pir4:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

Description		probable UV damage	UV-endonuclease ho	UV-endonuclease -	hypothetical profe	UV-endonuclease -	cell cycle protein	hypothetical prote	qlutamate synthase		probable ABC trans	hypothetical prote	_	Z				probable N-acetyim	probable FK506 oxi	conserved hypothet	N-acetyl-gamma-qlu	late expression fa	probable phosphotr	probable aldehyde	qlutaminefructos	rifamycin polyketi		C	probable multifunc	probable polyketid
SUMMARIES		/ 535	S55418	S55262	E83974	S71134	B82778	F72575	786785	386772	102583	T36004	175579	T17403	33118	137037	100052	175373	T17427	382965	:83562	110360	C83177	S27652	S01040	T17467	183253		528	H70984
DB	1		7	7			2							~													~		_	~
% Query Match Length I		320	320	929	322	599	437	313	1486	535	791	368	515	1809	1809	868	199	603	332	951	344	457	842	488	608	3413	319	390	397	1602
% Query Match		η,	6	16.7	16.4	15.1	6.7		6.5		•	6.2		6.2		•	6.1	•	0.9	•	٠	٠	•					•	•	
Score	, , ,	1320	303	255.5	250	230.5	102.5	98.5	98.5	96	96	95	95	94.5	4	94	93	66	92	92	91	90.5	90.5	90	90	90	δ	89.5	6	9.
Result No.	;	7	7	m	4	S	9	7	ω	σ	10	11	12	13	14	15	16	17	18	19	20	21	22	23	24	25	26	27	28	29

Page]		hypothetical prote	ABC-type transport	hypothetical prote	conserved hypothet	probable 2-compone	prolyl 4-hydroxyla	protéin F21D18.26	probable sun prote	TRK potassium upta	polyketide synthas	hypothetical prote	biosynthetic argin	probable serine/th	amyloid precursor-	conserved hypothet	tRNA nucleotidyltr	
	0	į	•															
7	Stower L.	H72516	T17482	T23472	A75252	G71511	149135	G96520	F75306	E84324	T30225	T33536	H83040	A70652	A46362	B75319	F81105	
	10	447 2	650 2	1221 2	253 2	352 2	537 2	707 2	439 2	558 2	10223 2	415 2	636 2	1110 1	653 2	360 2	417 2	
Z.		8.8	5.8	5.8	5.8	9. 2.8	5.8	5.8	5.7	5.7	5.7	5.7	5.6	5.6	5.6	9.5	5.5	
	~	68	88.5	88.5	88	88	88	88	87.5	87	87	86.5	98	86	85.5	82	84.5	
-39. rpr		30	31	32	33	34	35		37	38	39	40	41	42	43	44	45	

ALIGNMENTS

	probable UV damage endonuclease - Deinococcus radiodurans (strain R1)	
	endonuclease	
	V damage	
C75350	probable U	1

C:Species: Delinococcus andiodurans (Schaff K1)

R:White, O: / Eisen / J.A.; Heidelberg, J.F.; Hickey, E.K.; Peterson, J.D.; Dodson, R., M.; Shen, M.—Vamathevan, J.J.; Lam, P.; McDonald, L.; Utterback, T.; Zalewski, C. Science 286, 1571-1577, 1999

A:Title: Genome sequence of the radioresistant bacterium Deinococcus radiodurans R1.

A:Reference number: A75250; MUID:20036896

A:Status: preliminary

A; Molecule type: DNA

A; Residues: 1-326 <WHID
A; Cross-references: GB:AE002022; GB:AE000513; NID:g6459590; PIDN:AAF11370.1; PID:g645
A; Experimental source: strain R1
C; Genetics: A; Genetics: A;

0; Gaps); DB 2; Length 326; 2.3e-122; 1; Indels Score 1520; DB Pred. No. 2.3e-0; Mismatches 99.58; Query Match 99.5 Best Local Similarity 99.7 Matches 293; Conservative

.; 0

qq ò

61 YRLSSSLFPMLDLAGDDTGAAVLTHLAPQLLEAGHAFTDAGVRLLMHPEQFIVLNSDRPE 120 õ

121 VRESSVRAMSAHARVMDGLGLARTPWNLLLIHGGKGGRGAELAALIPDLPDPVRLRLGLE 180 QQ δy

q

181 NDERAYSPAELLPICEATGTPLVFDAHHHVVHDKLPDQEDPSVREWVLRARATWQPPEWQ 240 δλ

21.1 NDERAYSPAELLPICEATGTPLVFDAHHHVVHDKLPDQEDPSVREWVLRARATWQPPEWQ 270 241 VVHLSNGIEGPQDRRHSHLIADFPSAYADVPQIEVEAKGKEEAIAALRLMAPFK 294 qq ò

RESULT S55418

UV-endonuclease homolog ywjD - Bacillus subtilis C;Species: Bacillus subtilis

DB 2; Length 656;

16.7%; Score 255.5; DB 2; 31.8%; Pred. No. 6.6e-14; tive 38; Mismatches 103;

84; Conservative

```
167 ----PDLPDPVRLRLGLENDERAYSPAELLPICEATGTPLVFDAHHHVV------HDKL 215
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    404 KRNYARLSQSCKNRLVLENDDVGWTVHDLLPVCEELNIPMVLDYHHHNICFDPAHLREGT 463
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     56 HDIRLYRLSSSLFPMLDLAGDDTGAAVLTHLAPQLL-EAGHAFTDAGVRLLMHPEQFIVL 114
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              292 YGIRFLRLSSEMFP---FASHPVHGYKLAPFASEVLAEAGRVAAELGHRLTTHPGQFTQL 348
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            115 NSDRPEVRESSVRAMSAHARVMDGLGLARTFWN-----LLLLHGGKGGRGAELAALI--- 166
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             349 GSPRKEVVESAIRDLEYHDEL---LSLLKLPEQONRDAVMIIH--MGGQFGDKAATLERF 403
                                                                                                                                  A; Cross-references: EMBL: D11392; NID: 91526560; PID: 9927215
                        A;Status: nucleic acid sequence not shown
A;Molecule type: mRNA
A;Residues: 1-656 <YAJ>
                                                                                                                                                                                                                                                                                                                                                                           Best Local Similarity
                                                                                                                                                                                                                                                                                                                                      Query Match
                                                                                                                                                                               C; Genetics:
                                                                                                                                                                                                                                                                                                                                                                                                            Matches
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         RESULT.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 q
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    qq
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      δŏ
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           qq
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         Q
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       qq
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     Óλ
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              рp
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            δŏ
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               δ
C;Date: 15-Jul-1995 #sequence_revision 01-Sep-1995 #text_change 15-Oct-1999
R;Glaces. P.: Danchin, A.
Submitted to the EMBL Data Library, May 1995
A;Description: Cloning and sequencing of the Bacillus subtilis chromosomal region from 3
A;Reference number: $55418
A;Reference number: $60400
A;Reference number: $60580; MUID:98044033
A;Reference number: $60580; 
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                C; Species: Neurospora crassa
C;Date: 10-oct-1995 #sequence_revision 03-Nov-1995 #text_change 20-Jun-2000
C;Accession: S55562
B;Yajima, H.; Takao, M.; Yasuhira, S.; Zhao, J.H.; Ishii, C.; Inoue, H.; Yasui, A.
EMBO J. 14, 2393-2399, 1995
A;Title: A eukaryotic gene encoding an endonuclease that specifically repairs DNA damage
A;Reference number: S55262; MUID:95292980
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              73 LAGDDTGAAVLTHLAPQLLEAGHAFTDAGVRLLMHPEQFIVLNSDRPEVRESSVRAMSAH 132
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       81 VMWD-----FVTPEQKEFREIGELVKTHQLRTSFHPNQFTLFTSPKESVTKNAVTDMAYH 135
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             133 ARVMDGLGLA-RTPWNLLL--LHGGKGGRGAELAALIPDLPDPVRLRLGLENDERAYSPA 189
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    190 ELLPICEATGTPLVFDAHHHVVHDKLPDQEDPSVREWVLRARATW-----QPPEWQVVHL 244
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             196 ETLQVCEQEDVPFVFDFHHFYANP--DDHADLNVA--LPRMIKTWERIGLQPK----VHL 247
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     Gaps
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       16 RTVTLSRYRALSPAEREAKLLDLYSSNIKTLRGAADYCAAHDIRLYRLSSSLFPML.--D 72
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    21 KTLTFARYSKLSKTERKEALLTVTKANLRNTMRTLHYIIGHGIPLYRFSSSIVPLATHPD 80
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       245 SNGIEGPQDRRHSHLIADFPSAYADVPQIE-----VEAKGKEEAIAALRLM 290
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                248 SS----PKSEQAIRSHADYVDANFLLPLLERFRQWGTNIDFMIEAKQKDKAL--LRLM 299
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                19.8%; Score 303; DB 2; Length 320; 30.2%; Pred. No. 2.2e-18; Live 50; Mismatches 116; Indels 42;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            UV-endonuclease - Neurospora crassa
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    Best Local Similarity 30.2%
Matches 90; Conservative
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    Query Match
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              RESULT 3
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              δλ
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   q
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  Dp
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    οy
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   δ
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            qq
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           δ
```

```
RíTakami, H.; Nakasone, K.; Takaki, Y.; Maeno, G.; Sasaki, R.; Masui, N.; Fuji, F.; H
Nucleic Acids Res. 28, 4317-4331, 2000
A.Title: Complete genome Sequence for the alkaliphilic bacterium Bacillus halodurans a
A; Reference number: A83650; MUID:20263314
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     C.Species: Bacillus halodurans
C.Date: 01-Dec-2000 #sequence_revision 01-Dec-2000 #text_change 31-Dec-2000
                                                                                                                                                                                                                                                                                                                                                                                        hypothetical protein BH2597 (imported) - Bacillus halodurans (strain C-125)
216 PDQEDPSVREWVLRARATWQPPE-WQVVHLSNGIEG---PQDRR-HSHLIADFPSAYADV 270
                                                             464 LDISDPKLQE---RIANTWKRKGIKQKMHYSEPCDGAVTPRHRRKHRPRVMTLPPCPPDM 520
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       61 AYQISFFRLSSKLVPLVNHPLTEGWKYELA-----IAEELQAVGEFASEHQMRI 109
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   105 LMHPEOFIVLNSDRPEVRESSVRAMSAHARVMDGLGLARTPWNLLLLH--GGKGGRGAEL 162
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              163 AALIPD---LPDPVRLRLGLENDERAYSPAELLPICEATGTPLVFDAHHH-VVHDKLPDQ 218
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        168 EQFIENTASIPKSLLSMIMLENDDKSYTIDDVLYLGEKLAIPVVLDJHHHDVLH----- 221
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     110 DFHPDHFVVLNSEAKEITRRSLQTLLYHYKLLKGMEI--DPRHRCVLHVGGKKKGVEAGL 167
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     55 AHDIRLYRLSSSLFPML-----DLAGDDTGAAVLTHLAPQLLEAGHAFTDAGVRL 104
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      Gaps
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            4 QFGYVAMSMELANASPS----KTMTATQFEKTEDHEAGLRKLERIAKTNLHNCLRLLKHNL 60
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           1 QLGLVCLTV-----GPEVRFRTVTLSRYRALSPAEREAKLLD-LYSSNIKTLRGAADYCA 54
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  52;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         16.4%; Score 250; DB 2; Length 322; 26.2%; Pred. No. 7.6e-14; ive 68; Mismatches 116; Indels
                                                                                                                                          271 PQIEVEAKGKEEAIAALRLMAPFK 294
                                                                                                                                                                                                               521 -DLMIEAKDKEQAV--FELMRTFK 541
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  84; Conservative
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    Best Local Similarity
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                A; Accession: E83974
A; Status: preliminary
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               C; Accession: E83974
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 Query Match
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  Matches
```

```
C; Accession: B82778
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         q
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                ŏλ
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           qq
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    δ
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  q
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         ò
                                                                                                                                                                                                                                                                                                                C;Species: Schizosaccharomyces pombe
C;Species: Ja.Apr.1997 #sequence_revision 09-May-1997 #text_change 20-Jun-2000
C;Accession: S71134; T39815
R;Takao, M.; Yonemasu, R.; Yamamoto, K.; Yasui, A.
R;Takao, M.; Yonemasu, R.; Yamamoto, K.; Yasui, A.
R;Cles A.; Isologo, Inc. 1996
A;Title: Characterization of a UV endonuclease gene from the fisson yeast Schizosaccharc
A;Reference number: S71134; MUID:96188860
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          A;Status: preliminary; translated from GB/EMBL/DDBJ
A;Molecule type: DNA
A;Rosatidues: 1-599 CLINA
A;Rosatidues: 1-599 CLINA
A;Cross-references: EMBLAL023859; PIDN:CAA19577.1; GSPDB:GN00067; SPDB:SPBC19C7.09c
A;Experimental source: strain 972h-; cosmid c19C7
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        10;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       A; Cross-references: EMBL: D78571; NID: 91399000; PIDN: BAA11415.1; PID: 91399001
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        cell cycle protein XF0659 [imported] - Xylella fastidiosa (strain 9a5c)
C;Species: Xylella fastidiosa
C;Date: 18-Aug-2000 #sequence_revision 20-Aug-2000 #text_change 02-Sep-2000
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         249 RLGYACLNTILRSMKERVFCSRTCRITTIQRDGLESVKQLGTQNVLDLİKLVEWNHNFGI 308
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       119 PEVRESSVRAMSAHARVMDGLGLARTPWN------LLLLHGGKGGRGAE-----LAALI 166
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     36 EVVVDSAIRDLAYHDEI----LSRMKLNEQLNKDAVLIIHLGGTFEGKKETLDRFRKNY 420
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         167 PDLPDPVRLRLGLENDERAYSPAELLPICEATGTPLVFDAHHHVV-------HDKLPD 217
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       218 OEDPSVREWVLRARATWOPPEWOVVHLS----NGIEGPODRRHSHLJADFPSAYADVPQ 272
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              RLYRLSSSLFPMLDLAGDDTGAAVLTHLAPQLLEAGHAFTDAGVRLLMHPEQFIVLNSDR 118
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 309 HFMRVSSDLFP---FASHAKYGYTLEFAQSHLEEVGKLANKYNHRLTMHPGQYTQIASPR 365
  219 EDPSVREWVLRARATWQPPEWQV-VHLSNGIEGPQD-RRHSHLIADFPSAY----AD-V 270
                             Gaps
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         1 QLGLVCLTVGPEVRFRTVTLSRYRALSPAEREA--KLLDLYSSNIKTLRGAADYCAAHDI 58
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   A; Experimental source: strain SP972 R; Lyne, M.; Rajandream, M.A.; Barrell, B.G.; Lucas, M.; Gaillardin, C. submitted to the EMBL Data Library, June 1998
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       Length 599;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        Indels
                                                                                                                                                                                                                                                                                          UV-endonuclease - fission yeast (Schizosaccharomyces pombe)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           A,Status: preliminary, nucleic acid sequence not shown
A,Molecule type: mRNA
A,Residues: 1-599 <TAK>
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    DB 2;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     Mismatches 140;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    Score 230.5; DB Pred. No. 8e-12;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        42;
                                                                                                       271 POIEVEAKGKEEAIAALRLM 290
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    15.1%;
27.9%;
                                                                                                                                     : | : |:: :| :||
282 DHLHVMIEAKKKDLALFOLM 301
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        Conservative
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       273 IEVEAKGKEEAIAAL 287
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        : :||| ||:|: |
532 LMIEAKEKEQAVFEL 546
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    A; Reference number: Z21881
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             A; Gene: UVDE; SPBC19C7.09c
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    Query Match
Best Local Similarity
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      A; Accession: S71134
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        88;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                59
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        Matches
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     qq
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 qq
                                                   q
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         q
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            a
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                õ
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          ð
  ò
                                                                                                       δ
                                                                                                                                                        q
```

```
Risimpson, A.J.G.; Reinach, F.C.; Arruda, P.; Abreu, F.A.; Camargo, M.; Alvarenga, R. Briones, M.R.S.; Bueno, M.R.P.; Camargo, A.A.; Camargo, L.E.A.; Carraro, D.M.; Carreros, M.R.S.; Bueno, M.R.P.; Camargo, A.A.; Camargo, L.E.A.; Carraro, D.M.; Carreros, C.; Docena, C.; El-Dorry, H.; Facincani, A.P.; Ferreira, A.J.S.
submitted to Genbank, June 2000
A.Authors: Ferreira, V.C.A.; Ferro, J.A.; França, S.C.; França, S.C.; França, M.C.; Franco, M.C.; Fr
J.D.; Junqueira, M.L.; Kemper, E.L.; Kitajina, J.P.; Krieger, J.E.; Kuramae, E.E.; La
chado, M.A.; Martins, E.M.F.; Mastsukuma, A.Y.; Marcho, C.L.; Marques, M.V.; Martins
R.Authors: Martins, E.M.F.; Matsukuma, A.Y.; Menck, C.F.M.; Miracca, E.C.; Miyaki, C.
F.G.; Nunes, L.R.; Oliveira, M.C.; de Oliveira, R.C.; Palmieri,
Rodrigues, V.; Rosa, A.J.; de M.; de Rosa Jr., V.E.; de Sa, R.G.; Santelli, R.V.; Sawa
A.Authors: da Silva, A.C.R.; da Silva, A.M.; Silva Jr., W.A.; da Silva,
A.Rathako, M.H.; Vallada, H.; Van Sluys, M.A.; Verjovski-Almeida, S.; Vettore, A.L.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      A;Tille: Complete genome sequence of an aerobic hyper-thermophilic Crenarchaeon, Aero A;Reference number: A72450; MUID:99310339
                                                                                                                                                                                                                                                                        A;Molecule type: DNA
A;Residues: 1-437 <SIM>
A;Cross-references: GB:AE003910; GB:AE003849; NID:g9105532; PIDN:AAF83469.1; GSPDB:GN
               Seq
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         R;Kawarabayasi, Y.; Hino, Y.; Horikawa, H.; Yamazaki, S.; Haikawa, Y.; Jin-no, K.;
awa, H.; Takamiya, M.; Masuda, S.; Funahashi, T.; Tanaka, T.; Kudoh, Y.; Yamazaki,
DNA Res. 6, 83-101, 1999
R:anonymous, The Xylella fastidiosa Consortium of the Organization for Nucleotide Nature 406, 151-157, 2000
A:Title: The genome sequence of the plant pathogen Xylella fastidiosa.
A:Reference number: A82515; MUID:20365717
A:Note: for a complete list of authors see reference number A59328 below
A:Accession: B82778
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            12;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                A;Residues: 1-313 <KAW>
A;Cross-references: DDBJ:AP000062; NID:95105244; PIDN:BAA80891.1; PID:95105578
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    C.Species: Aeropyrum pernix
C.Date: 20-Aug-1999 #sequence_revision 20-Aug-1999 #text_change 20-Jun-2000
C.Accession: F72575
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    Gaps
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         98 IDAGVRLIMHPEQFIVLNSDRPEVRESSVRAMSAHARVMDGLGLARTPWNLLLLHGGKGG 157
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             158 RGAELAALIPDLPDPVRLRLGLE-NDERAYSPAE----LLPICEATGTPLVFDAHHHVV- 211
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  43 IKTLRGAADYCAAHDIRLYRLSSSLFPMLDLAGDDTGAAVLTHLAPQLLEAGHA----F 97
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   177 EDSSNADLRHDRNFLRIHV-LPLLHQRWPQATAVLAR---NAALAAANADLL-----
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                53;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      C;Superfamily: Aeropyrum pernix hypothetical protein APE1886
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             hypothetical protein APE1886 - Aeropyrum pernix (strain K1)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                212 HDKLPDQEDPS-----VREWVL-----RARATWQPPEWQ 240
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       84;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      DB 2;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  Score 102.5; DE Pred. No. 0.45;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                30; Mismatches
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   A;Gene: XF0659
C;Superfamily: hypothetical protein HI0404
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      6.78; 25.18;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          A, Experimental source: strain K1 C, Genetics:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  Query Match
Best Local Similarity 25.1%
Matches 56; Conservative
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             A; Contents: annotation
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              A; Status: preliminary
                                                                                                                                                                                                                                            A;Status: preliminary
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             A; Molecule type: DNA
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   A; Accession: F72575
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       A; Gene: APE1886
```

ĉ	90 00000	RESULT	6 1
W BK	Similarity 25:18; Pred. No. 0.64; 5; Conservative 35; Mismatches 84; Indels	Cab///	cab//2 conserved hypothe C;Species: Lactoo
Qy	11 PEVRFRTVTLSRYRALSP-AEREAKLLDLYSSNIKTLRGAADYCAAHDIRLYRLSSS 66	C;Date C;Acce R;Bold Genome	C;Date: 23-Mar-20 C;Accession: C867 R;Bolotin, A.; Wi Genome Res. in pr
Qy Db	67 LFPMLDLAGDDTGAAVLTHLAPQLLEAGHAFTDAGVRLLMHPEQFIVLNSDRPE 120 ::	A,Titl A,Refe A,Acce A,Stat	A;Title: The comp A;Reference numbe A;Accession: C867 A:Status: prelimi
Qy		A; Mole A; Resi A; Cros	ecule type: dues: 1-535 ss-reference
Qy	AALIPDLPDPVRLRLGLENDERAYSPAELLPICEATGTPLVEDAHHHVV	C; Gene A; Gene	C;Genetics: A;Gene: YljF
oy t	HDKLPDQEDPSVREWULRA 230	Quer Best Matc	Query Match Best Local Simi Matches 68;
a	ZOSMASVEEPSKRD-VIRS Z//	Qy	11 PEVRFRT
RESULT F86785	LT 8 8.5	qa	14 PVSKVKS
glut C;Sp	amate synthase (NADPH) (EC 1.4.1.13) large chain [imported] - Lactococcus lactis sub- ecies: Lactococcus lactis subsp. lartis	Qy	65 8
C; Da	te: 23-Mar-2001 #sequence_revision 23-Mar-2001 #text_change 31-Mar-2001 cession: F86785	q ₀	69 SKKIKNV
R; Bo Geno	R;Bolotin, A.; Wincker, P.; Mauger, S.; Jaillon, O.; Malarme, K.; Weissenbach, J.; Ehrli Genome Res. in press, 2001 A:Title: The complete genome sequence of the lastic acid hasterium	ठे त	
A; Re	ference number: A86625 cession: F86785	ag S	120 PIVLERG 165 LIPDLPD
A; MO A; MO	atus: preliminary lecule type: DNA	qa	: 11 169 RVRDL
A; Cr A; Ex	oss.references: GB:AE005176; NID:912724262; PIDN:AAK05384.1; GSPDB:GN00146 perimental source: strain IL1403	Oy	218
C; Ge	netics: ne: gltB	qq	214 QIIELGG
C; Ke	periamily: gintamare synthase (NADPH) ywords: oxidoreductase	RESULT T02583	10
Qu	%; Score 98.5; DB 2; Length 1486; 2%; Pred, No. 5;	probab N;Alte C;Spec	probable ABC tran N;Alternate names C;Species: Arabid
E .	; Conservative 33; Mismatches 81; Indels 105; Ga	C;Date C;Acce	C;Date: 05-Mar-19 C;Accession: T025
ž 2	14 KKYTYTLSKYRALIDLYSSARKEÄKLILIKEGAADYCAAHDTRIY	R;Roun submit A;Desc	R; Rounsley, S.D.; submitted to the A; Description: Ar
οy	62RLSSSLFPMLDLA-GDDTGAAVLTHLAPQLLEAGHAFTD 99	A;Refe A;Acce	A;Reference number A;Accession: T025
qa	288 CMEYFYRNGRELPESLLTMIPEAWGEOTGLSPELKAFYEYSTAHIAPWDGPAALVFTDGK 347	A; Molecule 1	A;Status: transta A;Molecule type: 1 A:Besidues: 1-701
Oy Dp	100 -AGVRLLMHPEQFIVLNSDRPEVRESSVRAMSAHARVMDGLGL 141 1	A; Expe	
Qy	ENDERAYSPAE-	M.; Ko euss,	M.; Koo, H.; Moffa euss, D.; Nierman
QQ	395PGNMLLVNTDEGKIIRNEEVKSYXANKYPYQEFLSAGLKKLSALTESEKTNPIPSAK 451	Nature 402, A;Title: Sec	402, 761-76 e: Sequence
oy :		A; Kele A; Acce A; Stat	A; Kelerence numbel A; Accession: C8481 A; Status: prelimir
gg ,	452 MNTLWKLFGYTDEVIRTVILPMAESANEPTISMGFDAPLAVLSDQ 496	A;Mole A;Resi A:Cros	A; Molecule type: I A; Residues: 1-791

```
betical protein yljF (imported) - Lactococcus lactis subsp. lactis (st
occcus lactis subsp. lactis
-001 #sequence_revision 23-Mar-2001 #text_change 23-Mar-2001
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            ; Kaul, S.; Lin, X.; Ketchum, K.A.; Crosby, M.L.; Brandon, R.C.; Syke

: EMBL Data Library, August 1998

- Trabidopsis thaliana chromosome II BAC T16B24 genomic sequence.

er: 214679
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          Tess: EMBL:AC004697; NID:g3402671; PID:g3402685
source: cultivar Columbia
S: Rounsley, S.D.; Shea, T.P.; Benito, M.I.; Town, C.D.; Fujii, C.Y
ffat, K.S.; Cronin, L.A.; Shen, M.; VanAken, S.E.; Umayam, L.; Tallon,
m. W.C.; White, O.; Eisen, J.A.; Salzberg, S.L.; Fraser, C.M.; Venter
768, 1999
                                                                                                                                  incker, P.; Mauger, S.; Jaillon, O.; Malarme, K.; Weissenbach, J.; Eh
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      e and analysis of chromosome 2 of the plant Arabidopsis thaliana.
er: A84420; MUID:20083487
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            17;
                                                                                                                                                                                                                                                                                                                                                                                                    es: GB:AE005176; NID:g12724145; PIDN:AAK05277.1; GSPDB:GN00146 source: strain 1L1403
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        A; Cross-references: GB: AE002093; NID: 93402685; PIDN: AAC28988.1; GSPDB: GN00139
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            999 #sequence_revision 05-Mar-1999 #text_change 16-Feb-2001
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         88; Gaps
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           -SLFPMLDLAGDDTGAAVLTHLAPQLLEAGHAFTDAGVRLLMHPEQFIVLNSDR 118
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 SVRAMSAHARVMDGLGLARTPWNLLLLHG-----GKGGRG----AELAA 164
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    G-QAVDERVKSIDKF-----LEGKLNPKSNVQFGEGGAGTFSDGKLTS 168
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         DPVRLRLGLENDERAYSPAELLPICEATGTPLVFDAHHHVVHDKLPD----- 217
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   ---RGRKVLEEFVQAGAPEDIL.-----YKAHPHVGTDLLRDIVKNIRK 213
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            TVTLSRYRALSPAEREAKLLD--LYSSNIKTL-RGAADYCAAHDIRL---YRLS 64
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               SLVLKKLRL----QESDLLDYRIYKESIDARHRGEIDFIYTVDIKLKDEARIL 68
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      unsporter At2g39190 [imported] - Arabidopsis thaliana
is: hypothetical protein T16B24.17
dopsis thaliana (mouse-ear cress)
                                                                                                                                                                     oress, 2001
plete genome sequence of the lactic acid bacterium.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      6.3%; Score 96; DB 2; Length 535; 24.4%; Pred. No. 2.1;
tive 36; Mismatches 87; Indels
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  GEVHFDAQVEEFLI-----EDEELQAVKLADG 246
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           ---QEDPSVREWVLRARATWQPPEWQVVHLSNG 247
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 ated from GB/EMBL/DDBJ
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            Conservative
                                                                                                                                                                                                                          ber: A86625
5772
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        1 <ROU>
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        1 <STO>
                                                                                                                                                                                                                                                                                                           inary
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            DNA
```

```
qq
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       a
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          οy
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        qq
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 q
                                                                                              qq
                                                                                                                                                                                            pp
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         à
                                                δy
                                                                                                                                                 δ
                                           A,Map position: 2
A;Introns: 110/1; 168/3; 216/1; 233/3; 281/3; 316/3; 358/3; 424/3; 460/2; 514/3; 536/3;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          R;Seeger, K.; Harris, D.; Bentley, S.D.; Parkhill, J.; Barrell, B.G.; Rajandream, M.A. submitted to the EMBL Data Library, July 1999
A;Reference number: 221574
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    A;Accession: T36004
A;Status: preliminary; translated from GB/EMBL/DDBJ
A;Mollcaule type: DDP
A;Residues: 1.368 <SEE>
A;Residues: 1.368 <SEE>
A;Cross-references: EMBL:AL096839; PIDN:CAB50757.1; GSPDB:GN00070; SCOEDB:SCC22.14c
                                                                                                                                                                                               18;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        Species: Streptomyces coelicolor
Date: 03-Dec-1999 #sequence_revision 03-Dec-1999 #text_change 08-Sep-2000
                                                                                                                                                                                                                                                                                                                                                                                                                                                                        281 QVYRGRTLDGADVAVKVQRPDLRHAVLRDIYILRLGLGVLRKVAKRENDIRVYADELGMG 340
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         66 PPARY----LARRRELGVMEADRRL--LRGSGITTYLVDAGLPGDLTGPTEMATAADADT 119
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          120 REIVRLELLAEQV-----ADTSGTVESFLA-NLAEAVHGAAVNAVAFTSVAGVRHGL 170
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           165 LIPDLPDPVRLRLGLENDERAYSPAELLPICEATGTPLVFDAHHHVVHDKLP-DQEDPSV 223
                                                                                                                                                                                                                                                                        167 LEVFSAFTIAAIRLRTSASVKGKSLEASEQIFGMVLKETMLHLGPTFIKVGQSLSTRPDI 226
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            158 -----ENDERAYS----- 187
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           188 -PAELLPICEATGTPLVFDAHHHVVHDKLPDQEDPSVREWVLRARATWQPPEWQVVHLSN 246
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         341 LAGELDFTLEAANASEFQEAHARFSYIRVPKVYQHLTRKRVLTM-----EWMVG---- 389
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       58 ---IRLYRLSSSLFPMLDLAGDDTGAAVLTHLAPQLLEAGH-----AFTD-AGVR--L 104
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        105 LMHPEQFIVLNSDRPEVRESSVRAMSAHARVMDGLGLARTPWNLLLLHGGKGGRGAELAA 164
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        171 ALAPEP-----BGPGEVRGAARWLT-----------------------GREVGG 197
                                                                                                                                                                                                                                                                                                                                           91 LEAGHAFTDAGVRL----SDRPEV 121
                                                                                                                                                                                                                                                                                                                                                                                                                                         122 --RESSVRAMSAHARVMDGLGLARTPW--NLLLLHGGKGG-------- 157
                                                                                                                                            Query Match 6.3%; Score 96; DB 2; Length 791;
Best Local Similarity 24.1%; Pred. No. 3.6;
Matches 83; Conservative 30; Mismatches 107; Indels 124; Gaps
                                                                                                                                                                                                                                            36 LDLYSSNIKTLRGAADYCAAHDIRLYRLSSSLFPMLDLAGDDTGAAVLTHLAP----QL 90
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                Gaps
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           --LRGAADYCAAHD--- 57
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 Superfamily: Streptomyces coelicolor hypothetical protein SCC22.14c
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        Ouery Match 6.2%; Score 95; DB 2; Length 368; Best Local Similarity 26.0%; Pred. No. 1.6; Matches 81; Conservative 27; Mismatches 103; Indels 100;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         247 GIEGPQDRRHSHLIADFPSAYADVPQIEVEAKGKEEAIAALRLM 290
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        hypothetical protein SCC22.14c - Streptomyces coelicolor
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           11 PEVRFRTVTLSRYRALSPAEREAKLLDLYSSNIKT-----
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  =
--
-
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            A, Experimental source: strain A3(2) C, Genetics:
                      A; Gene: At2g39190; T16B24.17
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            Gene: SCOEDB:SCC22.14c
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      Accession: T36004
C; Genetics
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   RESULT 11
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     QQ
                                                                                                                                                                                                                                               ò
                                                                                                                                                                                                                                                                                        g
                                                                                                                                                                                                                                                                                                                                             ò
                                                                                                                                                                                                                                                                                                                                                                                    QQ
                                                                                                                                                                                                                                                                                                                                                                                                                                         ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     qq
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           δ
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            Óγ
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       q
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          õ
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        q
```

```
A;Cross.references: GB:AE001863; GB:AE001825; NID:g6460670; PIDN:AAF12492.1; PID:g646
A;Experimental source: strain Rl
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           C;Accession: T17403
R;Reimmann, C.; Serino, L.; Beyeler, M.; Haas, D.
Microbiology 144, 3195-3148, 1998
A;Title: Dibydroaeruginoic acid synthetase and pyochelin synthetase, products of the A;Reference number: 218770; MUID:99061207
A;Accession: T17403
                                                                                                                                                                                                                                                                                                                                                                                                                                            hypothetical protein - Deinococcus radiodurans (strain R1)
(Species: Deinococcus radiodurans
(C.Species: Deinococcus radiodurans
(C.Date: 03-Dec-1999 #sequence_revision 03-Dec-1999 #text_change 31-Mar-2000
(C.Saccession: H75579
(R.White, O.; Eisen, J.A.; Heidelberg, J.F.; Hickey, E.K.; Peterson, J.D.; Dodson, M.; Shen, M.; Vamathevañ, J.J.; Lam, P.; McDonald, L.; Utterback, T.; Zalewski, S.; Smith, Ho.; Venter, J.C.; Fraser, C.M.
Science 286, 1571-1577, 1999
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             A;Title: Genome sequence of the radioresistant bacterium Deinococcus radiodurans A;Reference number: A75250; MUID:20036896
A;Accession: H75579
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            7,
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         A;Residues: 1-1809 CREI>
A;Cross-references: EMBL:AF074705; NID:93386352; PID:93386354; PIDN:AAC83657.1
C;Genetics:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       pyochelin synthetase (EC 6.3.2.-) - Pseudomonas aeruginosa
C.Species: Pseudomonas aeruginosa
C.Date: 15-Oct-1999 #sequence_revision 15-Oct-1999 #text_change 17-Nov-2000
                                                                                                               78 TGAAVLTHLAPQLLEAGHAFTDAGVRLLMHPEQFIVLNSDRPEVRESSVRAMSAHARVMD 137
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      138 GLGLARTPWNLLLL------HGGKGGRGAELAALIPDLPDPVRLRLGLENDERAYSPA 189
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       155 LLELAGEAWQLWQLARRQPRPAHSAQAGRARERVVLGGQVPDRVDWNRTLDLWGRGEFP- 213
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               ELLPICEATGTPLVFDAHHHVVHDKLPDQEDPSVREWVLRARATWQPPEWQVVHLSNGIE 249
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   ---HLLWIAVASGLPL--QLHAGLGEPGLRIDRTDPVL 239
                                                                    224 REWVLRARATWQPPEWQVVHLSNGIEGPQDRRHSHLIADFPSAYADVPQIEVEAKGKEEA 283
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  109 TAFAALARMVAELPE-----LQRRLHFE----GGDLPEALSGVQTLWTVPLRAQD 154
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                46;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        6.2%; Score 95; DB 2; Length 515; 24.5%; Pred. No. 2.5;
tive 16; Mismatches 86; Indels
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        A;Status: preliminary; translated from GB/EMBL/DDBJ
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            Best Local Similarity 24.59
Matches 48; Conservative
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               : || : :|:||
252 RDEVQRRFARALAEFP 267
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                250 GPQ-DRRHSHLIADFP 264
                                                                                                                                                                                                         284 IAAERL-MAPF 293
                                                                                                                                                                                                                                                                          294 VLAEILELAPF 304
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      1-515 <WHI>
198 ---ELSDPVLLR-
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         A;Status: preliminary
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             A; Molecule type: DNA
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   A; Map position: 2
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   A;Gene: DRA0265
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               Query Match
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         A;Gene: pchF
C;Function:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          A; Residues:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         RESULT 13
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               190
```

R1

```
Search completed: January 15, 2002, 13:52:39
Job time: 408 sec
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    Matches 51; Conservative
                                                                                                                                                                                                                                                                                                                                                                                                                                              1436 FGAGGDSLLIAQL 1448
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             225 EWVLRARATWOPPE 238
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  649 --IATAEARGEEPD 660
                                                                                                                                                                                                                                                                                                                                                                                      277 -- AKGKEEAIAAL 287
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               Query Match
Best Local Similarity
                                                                                                                                                  Qγ
                                                                                                                                                                                                                                                                                                                       QQ
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      δλ
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           qq
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              qq
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         QΥ
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  qq
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             Qγ
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      A.Title: Complete genome sequence of Pseudomonas aeruginosa PAO1, an opportunistic pathd A.Reference number: A82950; MUD:2043737
A.Reference number: A82950; MUD:2043737
A.Accession: C83118
A.Aoleoule type: DNA
A.Moleoule type: DNA
A.Moleoule type: DNA
A.Residues: 1-1809 <STO>
A.Cross-references: GB.AE004839; GB.AE004091; NID:99950434; PIDN:AAG07613.1; GSPDB:GN001
A.Experimental source: strain PAO1
C.Genetics: pchf; PA4225
C.Superfamily: acyl carrier protein homology; acetate--CoA ligase homology
C.Superfamily: acyl carrier protein homology <ACL>
F:560-1041/Domain: acetate--CoA ligase homology <ACL>
F:1410-1484/Domain: acyl carrier protein homology <ACC>
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   pyochelin synthetase PA4225 [imported] - Pseudomonas aeruginosa (strain PAO1)
C;Species: Pseudomonas aeruginosa
C;Date: 15-Sep-2000 #sequence_revision 15-Sep-2000 #text_change 31-Dec-2000
C;Accession: C83118
R:Stover, C.K.; Pham, X.Q.; Erwin, A.L.; Mizoguchi, S.D.; Warrener, P.; Hickey, M.J.; Bradman, S.; Yuan, Y.; Brody, L.L.; Coulter, S.N.; Folger, K.R.; Kas, A.; Larbig, K.; Lim, Nature 406, 959-964, 2000
A; Description: essential for the synthesis of pyochelin from dihydroaeruginoate
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        16;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   1208 FADQPWVRFGRFDMNGDLLDQGVAPHSVDILLSSG-ALNNALDTPALLAGLRELLSADAW 1266
                                                                                                                                                                                                                                                                                                                                                                               1155 PORRWRLLELGAGTAAASRTVIARLAPLVQ-----RGAEVDYLFT-DVSSYFLAAARER 1207
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      1267 LVIQELTREHNEISV----SQSLMMENPRDLRDERRQLFVHTGQWLEWLAAQGGDLACGV 1322
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                1155 PQRRWRLLELGAGTAAASRTVIARLAPLVQ-----RGAEVDYLFT-DVSSYFLAAARER 1207
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             1208 FADQPWVRFGRFDMNGDLLDQCVAPHSVDILLSSG-ALNNALDTPALLAGLRELLSADAW 1266
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              112 IVLNSDRPEVRESSVRAMSAHARVMDGLGLARTPWNLLLLHGGK-----GGRGAELA-AL 165
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              166 IP-----DLPDPVRLRLGLENDERAYSPAELLPICEATGTPLVFDAHHHVVHDKLPDQED 220
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 221 PSVREWVLRARATWQPPEWQVVHLSNGI-EGPQDRRHSHLIADFPSAYADVPQIEVE--- 276
                                                                                                                                                                                                                                                                                                                                                                                                                                           67 -----LFPMLDLAGDDTGAAVLTHLAPQLLEAGHAFTD-----AGVRLLMHPEQF 111
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 67 -----AGVRLLMHPEQF 111
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     Gaps
                                                                                                                                                                                                                                                                                                                          11 PEVRFRTVTLSRYRALSPAEREAKLLDLYSSNIKTLRGA-ADYCAAHDIRLYRLSSS--- 66
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          11 PEVRFRTVTLSRYRALSPAEREAKLLDLYSSNIKTLRGA-ADYCAAHDIRLYRLSSS--- 66
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       Ouery Match 6.2%; Score 94.5; DB 2; Length 1809; Best Local Similarity 24.9%; Pred. No. 14; Matches 78; Conservative 42; Mismatches 138; Indels 55;
                                                                                                                                                                                                                                                                  55;
                                                                                                                                                                                                       Query Match 6.2%; Score 94.5; DB 2; Length 1809; Best Local Similarity 24.9%; Pred. No. 14; Matches 78; Conservative 42; Mismatches 138; Indels 55;
                        A; Pathway: pyochelin synthesis
C; Superfamily: acyl carrier protein homology
C; Keywords: carrier protein; ligase
F;1410-1484/Domain: acyl carrier protein homology <ACP>
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             1436 FGAGGDSLLIAQL 1448
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  277 -- AKGKEEAIAAL 287
                                                                                                                                                                                                                                                                                                                          δλ
                                                                                                                                                                                                                                                                                                                                                                               q
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                q
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              ōλ
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   Db
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           οy
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 ŏ
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   q
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  δ
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          qq
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             δ
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      a
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 δλ
                                                                                                                                                                                                                                                                                                                                                                                                                                           δ
```

```
Cydrosesion: T37037

Rydrophy, L.; Harris, D.; Thomson, N.R.; Parkhill, J.; Barrell, B.G.; Rajandream, M.A subnited to the EMBL Data Library, August 1999
A; Reference number: 221619
A; Recession: T37037
A; Recent of the EMBL Data Library, August 1999
A; Recent of the EMBL Data Library, August 1999
A; Reference number: 221619
A; Recent of the EMBL Data Library, Lanslated from GB/EMBL/DDBJ
A; Residues: 1-898
A; Residues: 1-803
C; Genetics: Strain A3(2)
C; Genetics:
                                                1382 GKIDRKALTGFAR-OPQ----ADLRHGVAQAPADELENALLALWREV-LDNPSLGVEQDF 1435
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                pyruvate phosphate dikinase - Streptomyces coelicolor
C;Species: Streptomyces coelicolor
C;Dale: 03-Dec-1999 #sequence_revision 03-Dec-1999 #text_change 21-Jan-2000
                                                                                                                                                                                                                                            1323 VPPGSALDLLGYDVL.LARCKTDRARLEPAELLAFVEARVPRYMLPAQLRVL-ERLPVTGN 1381
112 IVLNSDRPEVRESSVRAMSAHARVMDGLGLARTPWNLLLLHGGK-----GGRGAELA-AL 165
                                                                                                                                                         1.66 IP-----DLPDPVRLRLGLENDERAYSPAELLPICEATGTPLVFDAHHHVVHDKLPDQED 220
                                                                                                                                                                                                                                                                                                                               221 PSVREWVLRARATWQPPEWQVVHLSNG1-EGPQDRRHSHLIADFPSAYADVPQIEVE--- 276
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          54; Gaps
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         73 LAGDDTGAAVLTHLAP-----QLLEAG-----HAFTDAGVRLLMHPEQF----IVLN 115
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    493 LSVDGTAGTVYAGAVPLVDSAVMRYLEAGAGGRATTGVVDAVARALARADAVRRLEVRAN 552
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           116 SDRPEVRESSVRAMSAHARVMDGLGLARTPWNLL------LLHGGKGGRGAELAAL 165
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             166 IPDLPDPVRLRLGLENDERAYSPAELLPICEA-TGTPLVFDAHHHVVHDKLPDQEDPSVR 224
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   6.2%; Score 94; DB 2; Length 898; 26.3%; Pred. No. 6.3; trive 22; Mismatches 67; Indels
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     A,Gene: SCOEDB:SCJ12.20
C;Superfamily: pyruvate,orthophosphate dikinase
```

*#				
			•	
			•	

```
January 15, 2002, 14:06:28; Search time 134.53 Seconds (without alignments) 630.625 Million cell updates/sec
                                                                                                                                                                                                                             US-09-724-296-38
3055
1 MGTTGLESLSLGDRGAAPTV.......HKLNKYHDWLWENHEKLSLS 580
GenCore version 4.5 Copyright (c) 1993 - 2000 Compugen Ltd.
                                                                                          OM protein - protein search, using sw model
                                                                                                                                                                                                                                Title:
Perfect score:
                                                                                                                                     Run on:
```

BLOSUM62 Gapop 10.0 , Gapext 0.5 Scoring table: Sequence:

Total number of hits satisfying chosen parameters: 473505 seqs, 146272329 residues Searched:

473505

Minimum DB seq length: 0 Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0% Maximum Match 100% Listing first 45 summaries

1: sp_archea:*
2: sp_bacteria:*
3: sp_fungi:*
4: sp_human:*
5: sp_invertebrate:*
6: sp_manmal:*
7: sp_mhc:*
8: sp_organelle:*
9: sp_phage:* sp_unclassified:* sp_plant:*
sp_rodent:*
sp_virus:*
sp_vertebrate:* SPTREMBL_17:* 10: Database :

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

	4 4 4 6 6 6 6 6 6 6 6 6 6 6 6 6 6 6 6 6	Description	095243 homo sapien	09z2d7 mus musculu	O9i9fl gallus gall	O9sfc1 arabidopsis		042403 gallus gall	09sf22 arabidopsis	025875 plasmodium	025730 plasmodium	026007 plasmodium	Ogu430 plasmodium	Q9u414 plasmodium	Q9u429 plasmodium	Q26104 plasmodium	Q9u0q6 plasmodium	O9u431 plasmodium	002061 caenorhabdi	Q9pum9 xenopus lae	043719 homo sapien
SUMMARIES	£	TT.	095243	Q922D7	Q919F1	Q9SFC1	Q9YGC6	042403	Q9SF22	025875	025730	026007	090430	090414	090429	026104	950060	090431	002061	6MD460	043719
	9	9 :	4	11	13	10	13	13	10	S	S	S	S	S	ഹ	ഹ	'n	ഹ	Ŋ	13	4
	Query	really car	580	554	416	419	467	344	702	782	782	782	782	782	782	782	2209	782	1359	282	755
æ	Query	יים יייי	100.0	59.6	28.8	11.3	7.2	9	4.8	4.7	4.6	4.6	4.6	4.6	4.5	4.5	4.5	4.5	4.5	4.4	4 . 4
	97.00	1000	3055	1821.5	880.5	344	220.5	202.5	146	143.5	141.5	141.5	140.5	140.5	138.5	137.5	137.5	136.5	136.5	135.5	135.5
	Result	2 :	1	7	æ	4	2	9	7	80	σ	10	11	12	13	14	15	16	17	18	19

Q9v433 drosophila		Q9ydp0 aeropyrum p	0921v0 helicobacte	Q9p91.6 pyrobaculum	Q9z2d8 mus musculu	095983 homo sapien	Q99730 homo sapien	Q9ubb5 homo sapien	09vzt4 drosophila	025262 helicobacte	09jmd6 mus musculu	Q9uhb7 homo sapien	Q9c9d8 arabidopsis	Q91fe4 arabidopsis	Q9u5e0 drosophila	Q91qv6 arabidopsis	09ayf2 oryza sativ	Q9jmx8 helicobacte	Q9w3dl drosophila	Q9pum8 xenopus lae	Q9c6s8 arabidopsis	Q9bkn8 plasmodium	Q9uis8 homo sapien	Q9wvc9 mus musculu
5 09V433 3 09USH9	5 Q9U7E0	1 Q9YDP0	2 Q9ZLV0	1 Q9P9L6	11 Q9Z2D8	4 095983	4 099730	4 Q9UBB5	5 Q9VZT4	2 025262	11 Q9JMD6	4 Q9UHB7		10 Q9LFE4	5 Q9USE0	10 Q9LQV6	10 Q9AYF2	2 09JMX8	5 Q9W3D1	13 Q9PUM8	10 090688	5 Q9BKN8	4 Q9UIS8	11 Q9WVC9
1058	1359	286	1819	230	285	291	754	411	537	1927	3122	1163	946	853	1058	648	970	1723	1183	303	561	881	302	873
4 4	4.4	4.3	4.3	4.3	4.3	4.3	4.3	4.3	4.3	4.3	4.2	4.2	4.2	4.2	4.2	4.2	4.2	4.2	4.1	4.1	4.1	4.1	4.1	4.1
135.5	133.5	132.5	132	131.5	131.5	131.5	131	130.5	130	130	129.5	129	128.5	1.28	127.5	127	127	127	126.5	126	126	126	125.5	125
20	22	23	24	25	56	27	28	29	30	31	32	33	34	35	36	37	38	39	40	41	42	43	44	45

ALIGNMENTS

```
Mamm.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              Q919F1
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         ID
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       QD
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  PT DT
  Qy
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              QY
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               q
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     δŏ
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   qq
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       δ
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      q
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         δò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        qq
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            ΟŊ
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             q
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 οy
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  qq
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        ŏ
                                                                                                                                                                                                                                                           0
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 SEQUENCE FROM N.A. MEDIATE-9845942; PubMed=9774669; MEDIATE-9845942; PubMed=9774669; Hendrich B., Bird A.; "Identification and characterization of a family of mammalian methyl-cpc binding proteins.";
                                                                                                                                                                                                                                                                                                                                                                                                                                            61 CNPLLOEPIASAQFGATAGTECRKSVPCGWBRVVKQRLFGKTAGRFDVYFISPQGLKFRS 120
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              121 KSSLANYLHKNGETSLKPEDFDFTVLSKRGIKSRYKDCSMAALTSHLONOSNNSNWNLRT 180
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         240
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       KYNKEALSPPRRKAFKKWTPPRSPFNLVQETLFHDPWKLLIATIFLNRTSGKMAIPVLWK 480
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 FLEKYPSAEVARTADWRDVSELLKPLGLYDLRAKTIVKFSDEYLTKQWKYPIELHGIGKY 540
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   121 KSSLANYLHKNGETSLKPEDFDFTVLSKRGIKSRYKDCSMAALTSHLONQSNNSNWNLRT 180
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          RSKCKKDVFMPPSSSSELQESRGLSNFTSTHLLLKEDEGVDDVNFRKVRKPKGKVTILKG 240
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             IPIKKTKKGCRKSCSGFVQSDSKRESVCNKADAESEPVAQKSQLDRTVCISDAGACGETL 300
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           241 IPIKKTKKGCRKSCSGFVQSDSKRESVCNKADAESEPVAQKSQLDRTVCISDAGACGETL 300
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               SVTSEENSLVKKKERSLSSGSNFCSEQKTSGIINKFCSAKDSEHNEKYEDTFLESEEIGT 360
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 KVEVVERKEHLHTDILKRGSEMDNNCSPTRKDFTGEKIFQEDTIPRTQIERRKTSLYFSS 420
                                                                                                                                                                                                                                                                                                                                                                                                                 CNPLLOEPIASAOFGATAGTECRKSVPCGWERVVKORLFGKTAGRFDVYFISPOGLKFRS 120
                                                                                                                                                                                                                                                           Gaps
                                                                                                                                                                                                                                                                                                            1 MGTTGLESLSLGDRGAAPTVTSSERLVPDPPNDLRKEDVAMELERVGEDEEQMMIKRSSE 60
                                                                                                                                                                                                                                                                                                                                       1 MGTTGLESLSLGDRGAAPTVISSERLVPDPPNDLRKEDVAMELERVGEDEEQMMIKRSSE 60
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           Mus musculus domesticus (western European house mouse).
Subaryota: Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Musinae; Mus.
NCBI_TaxID=10090, 10092;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   ..
                                                                                                                                                                                                         100.0%; Score 3055; DB 4; Length 580; 100.0%; Pred. No. 1.9e-219;
                                                                                                                                                                                                                                                           Indels
                                                                                                                             580 AA; 66050 MW; BF16FB21A34B8E5F CRC64;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      (TrEMBLrel. 10, Last sequence update) (TrEMBLrel. 17, Last annotation update)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      541 GNDSYRIFCVNEWKQVHPEDHKLNKYHDWLWENHEKLSLS 580
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           541 GNDSYRIFCVNEWKOVHPEDHKLNKYHDWLWENHEKLSLS 580
                                                                                                                                                                                                                                                                ..
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            554 AA.
                                                                                                                                                                                                                                                        0; Mismatches
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             01-MAY-1999 (TrEMBLrel. 10, Created)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         PRT;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           01-JUN-2001 (TrEMBLrel. 17, Last METHYL-CPG BINDING PROTEIN MBD4.
InterPro; IPR003265; Endo_3c.
Pfam; PF00730; HhH-GPD; 1.
Pfam; PF01429; MBD; 1.
SMART; SM00391; MBD; 1.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           Mus musculus (Mouse), and
                                                                                                                                                                                                                                  Best Local Similarity 100. Matches 580; Conservative
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         PRELIMINARY;
                                                                                                    Endonuclease
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      01-MAY-1999
                                                                                                                                SEQUENCE
                                                                                                                                                                                                         Query Match
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         0922D7
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           ~
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        181
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                301
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 361
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         181
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             241
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      421
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           RESULT
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              Q922D7
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               ID DATE OF THE SERVICE OF THE SERVIC
  DR DR DR DR SO
                                                                                                                                                                                                                                                                                                              δ
                                                                                                                                                                                                                                                                                                                                                               q
                                                                                                                                                                                                                                                                                                                                                                                                                 δ
                                                                                                                                                                                                                                                                                                                                                                                                                                                               q
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    a
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   Ġ
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      qq
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             q
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 δ
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   οy
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  q
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        δ
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       q
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         Q
```

```
63 PLLQEPIASAQFGATAGTECRKSVPCGWERVVKQRLFGKTAGRFDVYFISPQGLKFRSKS 122
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          51 SLLQEPTAST-LSSTTATEGHKPVPCGWERVVKQRLSGKTAGKFDVYFISPQGLKFRSKR 109
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             123 SLANYLHKNGETSLKPEDFDFTVLSKRGIKSRYKDCSMAALTSHLQNQSNNSNWNLRTRS 182
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      110 SLANYLLKNGETFLKPEDFNFTVLPKGSINPGYKHQSLAALTSLQPNETDVSKQNLKTRS 169
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               183 KCKKDVFNPPSSSSELQESRGLSNFTSTHLLLKEDEGVDDVNFRKVRKPKGKVTLLKGIP 242
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   243 IKKTKKGCRKSCSGFVQSDSKRESVCNKADAESEPVAQKSQLDRTVCISDAGACGETLSV 302
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            362
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           363 EVVERK--EHLHTDILKRGSEMDNNCSPTRKDFTGEKIFQEDTIPRTQIERRKTSLYFSS 420
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        394
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              421 KYNKEALSPPRRKAFKKWTPPRSPFNLVQETLFHDPWKLLIATIFLNRTSGKMAIPVLWK 480
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       481 FLEKYPSAEVARTADWRDVSELLKPLGLYDLRAKTIVKFSDEYLTKQWKYPIELHGIGKY 540
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          Gaps
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       6 LESLSLGD----RGAAPTVTSSERLVPDPPNDLRKEDVAMELERVGEDEEQMMIKRSSECN 62
                                                                                                                                            Hendrich B., Abbott C., McQueen H., Chambers D., Cross S., Bird A.; "Genomic structure and chromosomal mapping of the murine and human mbd1, mbd2, mbd3, and mbd4 genes.";
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        303 TSEENSLVKKKERSLSSGSNFCSEQKTSGIINKFCSAKDSEHNEKYEDTFLESEEIGTKV
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              339 --GDRKGEAHLHTGVLQDGSEMP-SCSQAKKHFTSE-TFQEDSIPRTQVEKRKTSLYFSS
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 1 MESPNLGDNRVRG-----ESLVPDPPWDRCKEDIAVGLGGVGEDGKDLVI--SSERS
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               230 SOKTKOKCRKSLLESTORNRKRASVVQKVGADRELVPQESQLNRTLCPADACA-RETVGL
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              455 FLEKYPSAEVARAADWRDVSELLKPLGLYDLRAKTIIKFSDEYLTKQWRYPIELHGIGKY
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          59.6%; Score 1821.5; DB 11; Length 554; 66.2%; Pred. No. 1.4e-127;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          49; Mismatches 116; Indels
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           N6_MTASE; UNKNOWN_1.
62577 MW; 792037CB180291F5 CRC64;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             541 GNDSYRIFCVNEWKQVHPEDHKLNKYHDWLWENHEKLSLS 580
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             515 GNDSYRIFCVNEWKQVHPEDHKLNKYHDWLWENHEKLSLS 554
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                01-0CT-2000 (TrEMBLrel. 15, Created)
01-0CT-2000 (TrEMBLrel. 15, Last sequence update)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             416 AA
Mol. Cell. Biol. 18:6538-6547(1998).
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             PRT;
                                                                                                                  MEDLINE=99373255; PubMed=10441743;
                                                                                                                                                                                                                                     Genome 10:906-912(1999).
                                                                                                                                                                                                                                                                  AF072249; AAC68878.1; -. AF120996; AAD56595.1; -.
                                                                                                                                                                                                                                                                                                                                                                                                                     InterPro; IPR002052; N6_Mtase.
Pfam; PF00730; HhH-GPD; 1.
Pfam; PF01429; MBD; 1.
                                                                                                                                                                                                                                                                                                                                                     InterPro, IPR003265; Endo_3c.
InterPro, IPR001739; MBD.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       Best Local Similarity 66.29 Matches 384; Conservative
                                                                                                                                                                                                                                                                                                                                  MGD; MGI:1333850; Mbd4.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           SMART; SM00391; MBD;
PROSITE; PS00092; N6__
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           554 AA;
                                                            SEQUENCE FROM N.A.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           SEQUENCE
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 Query Match
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         Q919F1;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          0919F1
                                                                                                                                                                                                                                                                        EMBL;
                                                                                                                                                                                                                                                                                                     EMBL;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                RESULT 3
```

11;

```
InterProj 1PR003265; Endo_3c. Pfam; PF00730; HiH-GPD; 1. SMART; SM00478: FNDO?c. 1
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           Nat. Genet. 19:187-191(1998)
EMBL; AC013483; AAF21203.1;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         Interpro; IPR001739; MBD.
Pfam; PF01429; MBD; 1.
SMART; SM00391; MBD; 1.
                                                                                                                         Conservative
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             PRELIMINARY;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    Xenopodinae; Xenopus.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    467 AA;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             SEQUENCE FROM N.A.
                                                   419 AA;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        SEQUENCE FROM N.A.
                                                                                                Query Match
Best Local Similarity
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        Kass S.U., Stroubd
"Xenopus MeCP2.";
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                NCBI_TaxID=8355;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     transcription.
                                                                                                                            .66
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     SEQUENCE
                                                    SEQUENCE
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             09YGC6;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              Q9YGC6
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  S
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                MECP2
                                                                                                                                                       205
                                                                                                                                                                                 88
                                                                                                                            Matches
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  RESULT
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           ΩD
                                                                                                                                                                                 qq
                                                                                                                                                                                                             δ
                                                                                                                                                                                                                                         gg
                                                                                                                                                                                                                                                                     δý
                                                                                                                                                                                                                                                                                              q
                                                                                                                                                                                                                                                                                                                          ó
                                                                                                                                                                                                                                                                                                                                                    qq
                                                                                                                                                                                                                                                                                                                                                                                  δy
                                                                                                                                                                                                                                                                                                                                                                                                           qq
                                                                                                                                                                                                                                                                                                                                                                                                                                        δy
                                                                                                                                                                                                                                                                                                                                                                                                                                                                   q
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                ò
  DR
DR
DR
SQ
                                                                                                                                                                                                                                                                                  6
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           Spermatophyta: Magnoliophyta; eudicotyledons; Core eudicots; Rosidae; eurosids II; Brassicales; Brassicaceae; Arabidopsis
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      374 DILKRGSEMDNNCSPTR------KDFTGEKIFQ--EDTIPRTQIERRKTSLY 417
                                                                                                                                                                                                                                                                                                              231 PKGKVTILKGIPIKKTKKGCRKSCSGFVQSDSKRESVCNKADAESEPVAQKSQLDRTVCI 290
                                                                                                                                                                                                                                                                                                                                                                     291 SDAGACGETLSVTSEENSLVKKKERSLSSGSNFCSEQKTSGIINKF-C-SAKDSEHNEKY 348
                                                                                                                                                                                                                                                                                                                                                                                              80 AEAGGRARRSSARSDAGSERGKRRRAAR-----GDREAAGVONEAPCEEAARVRHGQAV 133
                                                                                                                                                                                                                                                                                                                                                                                                                           ---HLHT 373
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 FSSKYNKEALSPPRRKAFKKWTPPRSPFNLVQETLFHDPWKLLIATIFLNRTSGKMAIPV 477
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                LWKFLEKYPSAEVARTADWRDVSELLKPLGLYDLRAKTIVKFSDEYLTKQWKYPIELHGI 537
                                Gallus gallus (Chicken).
Eukaryota, Metazoa, Chordata, Craniata, Vertebrata, Euteleostomi;
Archosauria, Aves, Neognathae, Galliformes, Phasianidae, Phasianinae;
                                                                                                                                                                                                                                                                                                                                      30 PDGSRAAPRDLPV-----RDGGEGGARSSQQRHGTAVRCE---RPSARGGKAERCATK 79
                                                                                                                                                                                                                                                                                     Gaps
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    STRAIN-CV. COLUMBIA;
Lin X., Kaul S., Town C.D., Benito M., Creasy T.H., Haas B.,
Ronning C.M., Koo H., Fujii C.Y., Utterback T.R., Barnstead M.E.,
Bowman C.L., White O., Nierman W.C., Fraser C.M.;
"Arabidopsis thaliana chromosome III BAC F17A17 genomic sequence.";
Submitted (DEC-1999) to the EMBL/GenBank/DDBJ databases.
                                                                                                                                                       "5-Methylcytosine DNA glycosylase activity is present in the human MBD4 (G/T mismatch glycosylase) and in a related avian sequence."; Submitted (APR-2000) to the EMBL/GenBank/DDBJ databases.
EMBL; AF257107; AAF68981.1;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  Eukaryota; Viridiplantae; Streptophyta; Embryophyta: Tracheophyta;
                                                                                                                                                                                                                                                                                                                                                                                                                                                        134 AELRAPPAAAGTQCVPAGQRPEPQPGAAADRLQGRPQSAAGGGERTGAGTAPAVLGEEES
                                                                                                                         Zhu B., Zheng Y., Angliker H., Schwarz S., Thiry S., Siegmann M., Jost J.-P.;
                                                                                                                                                                                                                                                                                     69;
                                                                                                                                                                                                                                                         DB 13; Length 416;
                                                                                                                                                                                                                                                                                    99; Indels
                    5-METHYLCYTOSINE G/T MISMATCH-SPECIFIC DNA GLYCOSYLASE.
                                                                                                                                                                                                      1107; AAF68981.1; -.
416 AA; 45454 MW; A3F70C6FF2133F2A CRC64;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     01-WAY-2000 (TrEMBLrel. 13, Created)
01-WAY-2000 (TrEMBLrel. 13, Last sequence update)
01-JUN-2001 (TrEMBLrel. 17, Last annotation update)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       538 GKYGNDSYRIFCVNEWKQVHPEDHKLNKYHDWLWENHEKLSL 579
     01-OCT-2000 (TrEMBLrel. 15, Last annotation update)
                                                                                                                                                                                                                                                        Query Match
28.8%; Score 880.5; DB 1:
Best Local Similarity 47.0%; Pred. No. 1.2e-57;
Matches 189; Conservative 45; Mismatches 99
                                                                                                                                                                                                                                                                                                                                                                                                                              EDTFLESEEIGTK-VEVVERKE--------
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        Arabidopsis thaliana (Mouse-ear cress).
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        PRT;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          PRELIMINARY;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          SEQUENCE FROM N.A.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            F17A17.27 PROTEIN.
                                                                                                                     SEQUENCE FROM N.A.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          NCBI_TaxID=3702;
                                                                                      NCBI_TaxID=9031;
                                                                                                                                                                                                                     SEQUENCE
                                                                                                                                                                                                                                                              Query Match
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       09SFC1;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       Q9SFC1
                                                                             Gallus
                                                                                                                                                                                                                                                                                                                                                                                                                                 349
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               RESULT
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             O9SFC1
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          ŏ
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     q
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        QΩ
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    δ
                                                                                                                                                                                                                                                                                                                                                                                                                                                          g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               ò
                                                                                                                                                                                                                                                                                                                                                                          ò
                                                                                                                                                                                                                                                                                                                                                                                                     g
                                                                                                                                                                                                                                                                                                                                                                                                                                 ð
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       ò
        ò
```

```
11;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           421 KYNKEALSPPRRKAFKKWTPPRSPFNLVQETLFHDPWKLLIATIFLNRTSGKMAIPVLWK 480
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      481 FLEKYPSAEVARTADWRDVSELLKPLGLYDLRAKTIVKFSDEYLTKQWKYPIELHGIGKY 540
                                                                                                                                                                                                                                                                141 ---DS--DSVCSKEGCSKVQAKVPRVSPYFQASTISQ--CDSDIVSSSOSGRNYRKGSSK 193
                                                                                                                                                                                                                                                                                                                                                                                                                                                                     306 ENSLVKK-----KERSLSSGSNFCSEQKTSGIINKFCSAKDSEHNEKYEDTFLESEEIGT 360
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                361 KVEVVERKEHLHTDILKRGSEMDNNCSPTRKDFTGEKIFQEDTIPRTQIERRKTSLYFSS 420
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            226 -- KVVKVSRYFHADGIQVNESQKEKSRNVRK-----TPIVSPVLSLSQKTDDVYL-- 273
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          274 ------RKTPDNTWVPPRSPCNLLQEDHWHDPWRVLVICMLLNKTSGAQTRGVISD 323
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              SNFTSTHLLLKEDEGVDDVNFRKVRKPKGKVTILKGIPI-----KKTKKGCRKSCSGF 257
                                                                                                                                                                                                                                                                                                                                             258 VQSDSKRESVCNKADAES-----EPVAQKSQLDRTVCISDAGACGET----LSVTSE 305
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 Enkaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Amphibia; Batrachia; Anura; Mesobatrachia; Pipoidea; Pipidae;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            Methylated DNA and MeCP2 recruit histone deacetylase to repress
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           MEDLINE-98282101; PubMed=9620779;
Jones P.L., Veenstra G.J.C., Wade P.A., Vermaak D., Kass S.U.,
Landsberger N., Strouboulis J., Wolffe A.P.;
                                                                                         11.3%; Score 344; DB 10; Length 419; 25.4%; Pred. No. 1.1e-17;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    Acustric (FEB-1998) to the EMBL/GenBank/DDBJ databases.
EMBL, AF106951; AAD03736.1; -.
EMBL, AF051768; AAD02651.1; -.
                                                                                                                                                     Indels
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               51757 MW; 5D3A719A59E560BC CRC64;
47957 MW; BAFF0BCA5A710C95 CRC64;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         OJINCO,
01.MAY-1999 (TrEMBLrel. 10, Created)
01.MAY-1999 (TrEMBLrel. 10, Last sequence update)
01.JUN-2001 (TrEMBLrel. 17, Last annotation update)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       194 ROVKVRRVSPYFQESTVSEQPN----QAPKGLRNYF----
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       467 AA
                                                                                                                                                            57; Mismatches
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        Strouboulis J., Wolffe A.P.;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               541 GNDSYRIFCVNEWKQVHPEDHKLNKYHDWL 570
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             Xenopus laevis (African clawed frog).
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                PRT;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 METHYL-CPG-BINDING PROTEIN MECP2.
```

```
---RKSV---
                                                                                                                                                            PRT;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                Matches 124; Conservative
                                                                                                                                                            PRELIMINARY;
71 SAQFGATAGTEC-----
                                                                                                                                                                                                                                                                                                                                                             COLUMBIA;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     Best Local Similarity
                                                                                                                                                                                                                                                                                                                                              SEQUENCE FROM N.A.
                                                                                                                                                                                                                               F26K24.5 PROTEIN.
                                                                                                                                                                                                                                                                                                                    NCBI_TaxID=3702;
                                                                                                                                                                                                                                                                                                                                                             STRAIN=CV.
                                                                                                                                                                                                                                              F26K24.5.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         Query Match
                                                                                                                                                         09SF22
                                                                                                                                           Q9SF22
                           qq
                                                        δò
                                                                                                                                                                        ŏ
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          QQ
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  qq
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             ŏλ
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        q
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 q
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       qq
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              δy
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      òγ
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              ŏλ
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         qd
                                       15;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     2
                                                                                                                                 100 -----TLPEGWIRKLKQFRSGRSAGKFDVYLINPGKAFRSKVELIAYFQKVGDISL 151
                                                                                                                  77 TAGTECRKSVPCGWERVVKQRLFGKTAGRFDVYFISPQGLKFRSKSSLANYLHKNGETSL 136
                                                                                                                                                                        137 KPEDFDFTVLSKRGIKSRYKDCSMAALTSHLQNQSNNSNWNLRTRSKCKKDV--FMPPSS 194
                                                                                                                                                                                                    152 DPNDFDFTV-TGRGSPSRREQ-----KQPKKPKAPKSSVSGRGRPKGSIKKVKPPVK 204
                                                                                                                                                                                                                                  SSELQESRGLSNFTSTHLLLK----EDEGVDDVNFRKV------RKPKGKVTILKGIP 242
                                                                                                                                                                                                                                                       205 SEGVQVKRVIEK-SPGKLLVKMPYSGTKEASDATTSQQVLVIKRGGRKRKSE-TDPSAAP 262
                                                                                                                                                                                                                                                                                       243 IKKTKKGCRKSCSGFVQSDSKRESVCNKADAESEPVAQKSQLDRTVCISDAGACGETLSV 302
                                                                                                                                                                                                                                                                                                                   263 KKRGRKPSNVSLAAAAAEAAKKKAI---KESSIKPLLE-----TVLPIKKRKTRETISV 313
                                                                                                                                                                                                                                                                                                                                               303 TSEE-----NSLVKK-----KERSLSSGSNFCSEQKTSGIINKFCSAKDSEHNEKYE 349
                                 60; Mismatches 157; Indels 67; Gaps
                                                                                                                                                                                                                                                                                                                                                                           314 DVKDTIKPEPLTPVIEKVMKGONPAKSPESRSTEGSPKIKTGLPKKELQQHHHHHHHH 373
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         Gallus gallus (Chicken).
Eukaryota: Metazoa; Chordata; Craniata; Vertebrata; Euteleostoml;
Archosauria; Aves; Neognathae; Galliformes; Phasianidae; Phasianinae;
                                                         22 SSERLVPDPPNDLRKEDVAMELER---VGEDEEQMMIKRS--SECNPLLQEPIASAQFGA 76
                                                                                      49 SSEH-OPGEPADEGKADMSESAEENLAVPESSASPKORRSVIRDRGPMYEDP----- 99
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   37; Gaps
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      16 AAPTVTSSERLVPDPPNDLRKEDVAMELER-----VGEDEEQMMIKRSSECNPLLQEPIA 70
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 5 AAAAAGGEERL----EEQADEGVAGLKERPPKAKKGRKERREDPEAEAEAEPSGAEPAE 59
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            MEDLINE-97415642; PubMed-9271441; Weitzel J.M., Buhrmester H., Straetling W.H.; "Chicken MAR-binding protein ARBP is homologous to rat methyl-CpG-
      7.2%; Score 220.5; DB 13; Length 467; 25.7%; Pred. No. 2e-08;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  6.6%; Score 202.5; DB 13; Length 344;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                57; Indels
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        344 344 346 MW; 53DD7BD9C9DF4FE3 CRC64;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   01-JAN-1998 (TrEMBLrel. 05, Created)
01-JAN-1998 (TrEMBLrel. 05, Last sequence update)
01-JUN-2001 (TrEMBLrel. 17, Last annotation update)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 01-JUN-2001 (TrEMBLrel: 17, Last annotation u
ATTACHMENT REGION BINDING PROTEIN (FRAGMENT).
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 Pred. No. 3e-07;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              344 AA.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         34.7%; Pred. No.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               PRT;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     binding protein MeCp2.";
Mol. Cell. Biol. 17:5656-5666(1997).
                                                                                                                                                                                                                                                                                                                                                                                                         350 DTFLESEEIGTKVEVVERKEHL 371
                                                                                                                                                                                                                                                                                                                                                                                                                                   374 HHHSESKASATSPEPETSKDNI 395
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   EMBL; Y14166; CAA74577.1; -.
InterPro; IPR000637; AT_hook.
InterPro; IPR001739; MBD.
                                 Conservative
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            PRELIMINARY;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             Conservative
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           PRINTS; PR00929; ATHOOK.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              Pfam; PF01429; MBD; 1
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          SMART; SM00391; MBD;
NON_TER 344 3
                   Local Similarity
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              Local Similarity
des 59; Conserv
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            SEQUENCE FROM N.A.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    NCBI_TaxID=9031;
       Query Match
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        NON_TER
SEQUENCE
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                Query Match
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          042403
                                                                                                                                                                                                                                  195
                               Matches
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          Matches
                                                                                                                                                                                                                                                                                                                                                                                                                                                                              RESULT
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             042403
                                                                                      qq
                                                                                                                  ò
                                                           δ
                                                                                                                                                                        δý
                                                                                                                                                                                                    qq
                                                                                                                                                                                                                                                            QQ
                                                                                                                                                                                                                                  δ
                                                                                                                                                                                                                                                                                         ò
                                                                                                                                                                                                                                                                                                                    Dp
                                                                                                                                                                                                                                                                                                                                                 ολ
                                                                                                                                                                                                                                                                                                                                                                           qq
                                                                                                                                                                                                                                                                                                                                                                                                                                   g
                                                                                                                                                                                                                                                                                                                                                                                                         ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 g
```

```
27;
                                                 60 AGKADGSGGTAAAPAVPEASASPKQRRSIIRDRGPMYDDPTLPEGWTRKLKQRKSGRSAG 119
----PCGWERVVKQRLFGKTAG 104
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               Arabidopsis thaliana (Mouse-ear cress).
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta; Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; Rosidae; eurosids II; Brassicales; Brassicaceae; Arabidopsis.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   87 PCGWE-----RVVKQRLFGKTAGRFDVYFISPQGLKFRSKSSLANYLHKNGETSL 136
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 137 KPEDFDFTVLSKRGIKSRYKDCSMAALTSHLQ-----NQSNNSNWNL-RTRSKCKKD- 187
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       101 YPWEITFSVFT-NGMKQCQKNKAPVVGTAFLNLAEYACVTDKKEFDINIPLILSACVASE 159
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                188 ----VFM------PPSSSSELQ------ESRGLSNFTSTHLLLKEDEGVDDVNF 225
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     160 THPLLFVSLSLLELRTTPETSDSAAQTAVVPLPLPSPSPQQPTETHSVEKEDVSAIKAGL 219
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          226 RKVRKPKGKVTILKGIPIKKTKKGCRKS---CSGFVQSDSKRESVCNKADAESEPVAQKS 282
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   220 RKVK-----IFTEFVSTRKAKKACREEEGRFSSFESSESLDDFETDFDEGKEELMSMRK 273
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      80; Mismatches 226; Indels 170; Gaps
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     28 PDPPNDLRKEDVAMELERV-GEDEEQMMIKRSSECNPLLQEPIASAQFGATAGTECRKSV 86
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       339 AKDSEHNEKYEDTFLESEEIGTKVEVVERKEHLHTDILKRGSEMDNNCSPTRKDFTGEKI 398
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            320 AEDSAAGLVYEASLLPRRSI-----LPWRKRKLSFRSPKSKGE-----PLLKKDNGEEG 368
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     399 FQEDTIPRTQIERRKTSLYFSSKYNKEALSPPRRK------AFKKWTPPRSPFNLVQ 449
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          369 GDDIDFDRRQLSSDEAHPPFGSKIDEDSSANPRTSFSEFGEDSFAIGSWEE-----K 420
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               450 ETLEHDPWKLLIATIFL-----NRTSGKMAIPVLWKFLEKYPSAEVARTADWRDVSEL 502
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          10 PWPPLVTRKYEVKLSVKKLEGWD------LVREGVPEKD-----RLTV 46
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    Lin X., Kaul S., Town C.D., Benito M., Creasy T.H., Haas B.,
Ronning C.M., Koo H., Fujii C.Y., Utterback T.R., Barnstead M.E.,
Bowman C.L., White O., Nierman W.C., Fraser C.M.;
Arabidopsis thaliana chromosome III BAC F26K24 genomic sequence.";
Submitted (JAN-2000) to the EMBL/GenBank/DDBJ databases.

EMBL: AC016795; AAF23192.1;
SEQUENCE 702 AA; 79173 MW; 6F2EDC7F8ED1F038 CKC64;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                283 QLDRTVCISDAGACGETLS----VTSEENSLVKKKERSLSSGSNFCSEQKTSGIINKFCS
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  4.8%; Score 146; DB 10; Length 702; 20.7%; Pred. No. 0.012;
                                                                                                                       105 RFDVYFISPQGLKFRSKSSLANYLHKNGETSLKPEDFDFTVLSKRGIKSR 154
                                                                                                                                                                 01-MAY-2000 (TrEMBLrel. 13, Created)
01-MAY-2000 (TrEMBLrel. 13, Last sequence update)
01-MAY-2000 (TrEMBLrel. 13, Last annotation update)
                                                                                                                                                                                                                                                                                                                                                                                        702 AA
```

) Db 421 EV		Qy 489EVARTADWRDVSELLKPLGLYDL
Qy 503 LKPI 	503 LKPLGLYDLRAKTIVKESDEYLTKQWKYPIELHGIGKYGNDSYRIFCVNE-WKQVHPEDH 561 	535 411
SUL 587	PRELIMINARY; PRT; 782 AA.	
DT 01-NOV-DT 01-JUN-DT 01-	01-NOV-1996 (TrEMBLrel. 01, Created) 01-NOV-1996 (TrEMBLrel. 01, Last sequence update) 01-JUN-2000 (TrEMBLrel. 14, Last annotation update) PR86 RHOPTRY PRECURSOR PROTEIN.	1D 025/30 FKELIMINAKI; PKT; /82 AC 025/30; PKT: /82 AC 025/30; DT 01-NOV-1996 (TrEMBLrel. 01, Last sequence DT 01-MAR-2001 (TrEMBLrel. 16, Last annotati
	Plasmodium falciparum. Eukaryota; Alveolata; Apicomplexa; Haemosporida; Plasmodium. NCBL_TaxID=5833;	DE RHOPTRY ASSOCIATED PROTEIN-1. GN RAP-1. OS Plasmodium falciparum (isolate FC27 / Pap OC Eukaryota; Alveolata; Apicomplexa; Haemos
	SEQUENCE FROM N.A. STRAIN-HONDURAS I CDC; MEDLINE-90348711; PubMed-2200961; RIGHS R.G., Takacs B., Lahm H.W., Delives C.J., Goman M., Certa V., Matile H., Woollett G.R., Scaife J.G.;	
	"Characterisation and sequence of a protective rhoptry antigen from Plasmodium falciparum."; Dals Biochem. Parasitol. 41:125-134(1990).	DR EMBL; U20985; AAA63681.1; SQ SEQUENCE 782 AA; 90023 MW; F69E26A2A
	SEQUENCE FROM N.A. STRAIN-HONDURAS I CDC; MEDLINE-92244303; Pubmed-1574089;	Query Match 4.6%; Score 141.5; Best Local Similarity 21.1%; Pred. No. 0. Matches 110; Conservative 71; Mismatche
	HTMWATOR.F.; #THP Sequence of the p82 rhoptry protein is highly conserved between two Plasmodium falciparum isolates."; #WOL Bicochem. Parasitol. 51:327-330(1992).	Qy 123 SLANYLHKNGETSLKPEDFDFTVLSK :: :: ::
SQ SEQUENC	2000/); AAAZ AA; 90096 MW; DIAD099862528D42 CRC64;	172 -NNSNWNLRTRSKCKKDVFMPP ::
Query Match Best Local Matches 11:	ch il Similarity 21.5%; Pred. No. 0.022; 112; Conservative 70; Mismatches 175; Indels 165; Gaps 26;	76
Oy 123 SI :: Db 18 NV	23 SLANYLHKNGETSLKPEDFDFTVLSKRGIKSRYKDCSMAALTSHLONQS-171 :: :: : : : : : 13 18 NVADGINVNGDNNYGKTIINNDFNFDDYNYWTPINKKEFLNSYEDKFSSESFLENKSS 75	123
Qy 172 -N : Db 76 VE	72 -NNSNWNLRTRSKCKKDVFMPPSSSSELQESRGLSNFTSTHLLLKEDEGVDDVNFRKVRK 230 :	Db 159 KSASPHGESNSSEESTTKSSKRSAS Qy 351 TFLESEEI-GTKVEVVERKEHLHTDILKRGS
Qy 231 PK Db 123 -k	RTVCI SNV	Db 201 TLTPLEELYPTNVNLFNYKYSLNNMEENINILKNEC Qy 396EKIFQEDTIPRTQIERRKTSLYFS
Qy 291 SE Db 159 KS	SDAGACGETLSVTSEENSLVKKKERSLSSGSNFCSEQKTSGIINKFCSAKDSEHNEKYED 350	Db 261 ALEKIGKESDEAPFMFSENKFLENOVKERNVAGSFS OY 444 PFNLVQETLEHDPWKLLIATIFLNRTSGKMAIPVLK
Qy 351 TF Db 201 TE	TFLESEEI-GTKVEVVERKEHLHTDILKRGSEMDNNCSPTRKDFTG 395 : 3:1 TLTPLEELYPINVNLFNYKYSLNNMEENINILKNEGDLVAOKEEFEYDENMEKAKODKKK 260	Db 308 -DEVIEKTEVSKKTFSGIGFNLTDKEAKVLGVG Qy 489EVARTADWRDVSEILKPLGLYDL
Qy 396 Db 261 AI	396EKIFQEDTIPRTQIERRKTSLYFSSKYNKEALSPPRKKAFKKWTPPRS 443	Db 364 ESLQGRIID1KKRESMISTTFEQQKECLKNMGVLDI Qy 535 HGIGKYGNDSYRIF-CVNEWKQVHPEDF
Oy 444 PF Db 308 -L	444 PFNLVQETLFHDPWKLLIATIFLNRTSGKMAIPVLWKFLEKYPSA	Db 411 -CIGSFGEHHLRLYEFENDLFKFHPNIDYLTLADGY

```
25;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              NFTSTHLLLKEDEGVDDVNFRKVRK 230
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              KRGIKSRYKDCSMAALTSHLQNOS- 171
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                | :|:||: ||: || RSASAAILEEDDSKDDMEF---- 122
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    SVCNKADAESEPVAQKSQLDRTVCI 290
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 SPSN----- 158
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           EQKTSGIINKFCSAKDSEHNEKYED 350
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                | | | ::|
GDLVAQKEEFEYDENMEKAKQDKKK 260
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      SSKYNKEALSPPRRKAFKKWTPPRS 443
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           WKFLEKYPSA----- 488
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                ::||
3ATYQEYPETMLYNCPNNSNLFDTI 363
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  LRAK-TIVKFSDEYLTKQWKYPIEL 534
LRAK-TIVKFSDEYLTKQWKYPIEL 534
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  S-----EMDNNCSPTRKDFTG 395
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       | ::| |:|
SRFFSK--LNP----FKK----- 307
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           ; DB 5; Length 782;
.03;
es 176; Indels 165;
                                                                                                                                                                                                                                                pua New Guinea).
sporida; Plasmodium.
                     LELNDTQCKFGT----
                                                                                                                                                                                                                                                                                                                                                 /DDBJ databases.
                                                                                                                                                                                                                                                                                                                                                                           564C8EA CRC64;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        : II
LELNDTQCKFGT--
                                                    )HKLNKYH 567
|-|| | |
|XKLQKNH 451
                                                                                                                                                                                             se update)
tion update)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        HKLNKYH 567
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         :(| | | 4
YKLQKNH 451
                                                                                                                                                      AA.
```

```
STRAIN-FUP;
                                                                                                                                                                                                                SEQUENCE
                                                                                                                         Saul A.;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  090414;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  090414
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    RESULT 12
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   δ
                                                                                                                                                                                                                                                                                                                                                                                                                                                                         qq
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       δ
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    QQ
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  δ
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              Dp
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             Οy
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        QQ
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        9
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 ōλ
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       q
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                231 PKGKVTILKGIPIKKTKKGCRKSCSGFVQSDSKRESVCNKADAESEPVAQKSQLDRTVCI 290
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     123 -KASPSVVKTSTPSGTQTSGLKSSS---PSSTKSSSPSN-------V 158
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 172 - NNSNWNLRTRSKCKKDVFMPPSSSSELQESRGLSNFTSTHLLLKEDEGVDDVNFRKVRK 230
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  76 VDDGNINLTDTSTSNK-----SKKGHGRSRVRSASAAAILEEDDSKDDMEF---- 122
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           291 SDAGACGETLSVTSEENSLVKKKERSLSSGSNFCSEQKTSGIINKFCSAKDSEHNEKYED 350
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      159 KSASPHGES---NSSEESTTKSSKRSAS-----VAGIV----GADEEAPPAPKN 200
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      351 TFLESEEI-GTKVEVVERKEHLHT-----DILKRGS------EMDNNCSPTRKDFTG 395
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 201 TLTPLEELYPTNVNLFNYKYSLNNMEENINILKNEGDLVAQKEEFEYDENMEKAKQDKKK 260
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               396 --EKI------FQEDTIPRTQIERRKTSLYFSSKYNKEALSPPRRKAFKKWTPPRS 443
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          444 PFNLVQETLFHDPWKLLIATIFLNRTSGKMAIPVLWKFLEKYPSA------- 488
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       308 -DEVIEKT---EVSKKTFSGIGFNLTDKEAKVLGVGATYQEYPETMLYNCPNNSNLFDTI 363
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    ---EVARTADWRDVSELLKPLGLYDLRAK-TIVKFSDEYLTKQWKYPIEL 534
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  364 ESLQGRIIDIKKRESMISTTFEQQKECLKNMGVLDLELNDTQCKFGT------410
                                                                                                                                                                                                                                                                                                                                                                                                                                          123 SLANYLHKNGET----SLKPEDFDF----TVLSKRGIKSRYKDCSMAALTSHLQNQS- 171
                                                                                                                                                                                                                                                                                                                                                                                                                 Gaps
                                                                                                                                                                                                                                                                                                                                                                                                                                                                       18 NVADGINVNGDNNYGKTIINNDFNFDDYNYWTPINKKEFLNSYED--EFSSESFLENKSS 75
                                                                                                                                                                               Ridley R.G., Takacs B., Lahm H.W., Delves C.J., Goman M., Certa V., Matile H., Woollett G.R., Scaife J.G.; "Characterisation and sequence of a protective rhoptry antigen from Plasmodium falciparum."; Plasmodium falciparum."; Mol. Biochem. Parasitol. 41:125-134(1990).

EMBL; M32853; AAA29753.1; -.
                                                                                                                                                                                                                                                                                                                                                                                            Best Local Similarity 21.18; Pred. No. 0.03;
Matches 110; Conservative 71; Mismatches 176; Indels 165;
                                                                                                                                                                                                                                                                                                                                                                                 4.6%; Score 141.5; DB 5; Length 782;
                                                                                                         Alveolata; Apicomplexa; Haemosporida; Plasmodium
                                           01-NOV-1996 (TrEMBLrel. 01, Last sequence update)
01-JUN-2000 (TrEMBLrel. 14, Last annotation update)
RHOPTRY ASSOCIATED PROTEIN 1 (ROP1) PRECURSOR (FRAGMENT).
                                                                                                                                                                                                                                                                                                                                    782 AA; 90080 MW; 2F2E240D40A4C902 CRC64;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               535 HGIGKYGNDSYRIF-CVNEWKQVHP-----EDHKLNKYH 567
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  411 -CIGSFGEHHLRLYEFENDLFKFHPNIDYLTLADGYKLOKNH 451
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 Last sequence update)
Last annotation update)
   782 AA
                                                                                                                                                                                                                                                                                           POTENTIAL
                                                                                                                                                                                                                                                                                                         POTENTIAL
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                01-MAY-2000 (TrEMBLrel. 13, Created)
01-MAY-2000 (TrEMBLrel. 13, Last sequ
01-MAY-2000 (TrEMBLrel. 13, Last anno
                              01-NOV-1996 (TrEMBLrel. 01, Created)
01-NOV-1996 (TrEMBLrel. 01, Last seq
                                                                                                                                                    SEQUENCE FROM N.A.
MEDLINE=90348711; PubMed=2200961;
PRELIMINARY;
                                                                                                                                                                                                                                                                                      1 22
23 >782
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      PRELIMINARY;
                                                                                         Plasmodium falciparum.
                                                                                                                     NCBI_TaxID=5833;
                                                                                                         Eukaryota;
                                                                                                                                                                                                                                                                                                                    NON_TER
SEQUENCE
                                                                                                                                                                                                                                                                                                                                                                                   Query Match
                                                                                                                                                                                                                                                                        Signal.
SIGNAL
              026007;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   Q9U430;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    090430
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        RESULT 11
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        090430
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      οy
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       Qγ
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          δ
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               δy
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          qq
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         ογ
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     В
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     Qγ
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             à
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           엄
```

```
"Efficacy of vaccines containing Rhoptry-Associated Proteins RAP1 and RAP2 of Plasmodium falciparum in Saimiri boliviensis monkeys."; Submitted (Nov-1999) to the EMBL/GenBank/DDBJ databases. EMBL; AF205283; AA:23404.1; -.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     Query Match 4.6%; Score 140.5; DB 5; Length 782;
Best Local Similarity 21.5%; Pred. No. 0.036;
Matches 112; Conservative 70; Mismatches 175; Indels 165; Gaps
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  172 -NNSNWNLRTRSKCKKDVFMPPSSSSELQESRGLSNFTSTHLLLKEDEGVDDVNFRKVRK 230
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   76 VDDGNINLTDTSTSNK-----SKKGHGRSRVRSASAAAILEEDDSKDDMEF---- 122
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      231 PKGKVTILKGIPIKKTKKGCRKSCSGFVQSDSKRESVCNKADAESEPVAQKSQLDRTVCI 290
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      123 -KASPSVVK---------TSTPSCTQTSGSK-----SSSPSSTKSSSPSNV-- 158
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           291 SDAGACGETLSVTSEENSLVKKKERSLSSGSNFCSEQKTSGIINKFCSAKDSEHNEKYED 350
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               159 KSASPHGES---NSSEESTTKSSKRSAS-----VAGIV----GADEEAPPAPKN 200
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               201 TLTPLEELYPTNVNLFNYKYSLNNMEENINILKNEGDLVAQKEEFEYDENMEKAKQDKKK 260
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    396 --EKI------FQEDTIPRTQIERRKTSLYFSSKYNKEALSPPRRAFKKWTPPRS 443
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    307
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    308 - DEVIEKT - - - EVSKKTFSGIGFNLTYKEAKVLGVGVTYQEYPETMLYNCPNNSNLFDTI 363
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        364 ESLOGRIIDIKKRESMISTTFEQQKECLKNMGVLDLELNDTQCKFGT------- 410
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               351 TFLESEEI-GTKVEVVERKEHLHT-----DILKRGS------EMDNNCSPTRKDFTG 395
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        --EVARTADWRDVSELLKPLGLYDLRAK-TIVKFSDEYLTKQWKYPIEL 534
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              123 SLANYLHKNGET----SLKPEDFDF-----TVLSKRGIKSRYKDCSMAALTSHLQNQS- 171
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       18 NVADGINVNGDNNYGKTIINNDFNFDDYNYWTPINKRFFLNSYED--EFSSESFLENKSS 75
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               261 ALEKICKESDEEPFMFSENKFLENQVKERNVAGSFSRFFSK--LNP----FKK----
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        444 PFNLVQETLFHDPWKLLIATIFLNRTSGKMAIPVLWKFLEKYPSA-------
                                                                                                               Eukaryota; Alveolata; Apicomplexa; Haemosporida; Plasmodium.
NCBL_TaxID=58%;;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         Eukaryota; Alveolata; Apicomplexa; Haemosporida; Plasmodium.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         782 782
782 AA; 90131 MW; A0FCB64529C34E59 CRC64;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             535 HGICKYGNDSYRIF-CVNEWKQVHP-----EDHKLNKYH 567
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             411 -CIGSFGEHHLRLYEFENDLFKFHPNIDYLTLADGYKLQKNH 451
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     01-MAY-2000 (TrEMBLrel. 13, Created)
01-MAY-2000 (TrEMBLrel. 13, Last sequence update)
01-MAY-2000 (TrEMBLrel. 13, Last annotation update)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         782 AA
RHOPTRY-ASSOCIATED PROTEIN 1 (FRAGMENT)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     PRT;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 RHOPTRY ASSOCIATED PROTEIN-1.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         PRELIMINARY;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     Plasmodium falciparum.
                                                                               Plasmodium falciparum.
                                                                                                                                                                                                                                                        SEQUENCE FROM N.A.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     SEQUENCE FROM N.A. STRAIN=FCC1/HN;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          NCBI_TaxID=5833;
```

26;

```
Query Match
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              026104
                                                                                                                                                                                                                                                                                                                                                     396
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            qq
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         qq
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  Dp
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          qq
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             δŏ
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           qq
                                                                                                                                                                                                                                                                                                                   a
                                                                                                                                                                                                                                                                                                                                                                                                                        ŏ
                                            q
                                                                                 ŏ
                                                                                                               QQ
                                                                                                                                                  ŏ
                                                                                                                                                                                q
                                                                                                                                                                                                                  ò
                                                                                                                                                                                                                                                  qq
                                                                                                                                                                                                                                                                                  ò
                                                                                                                                                                                                                                                                                                                                                     ò
                                                                                                                                  25;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       "Efficacy of vaccines containing Rhoptry-Associated Proteins RAP1 and RAP2. of Plasmodium falciparum in Saimiri boliviensis monkeys."; ...iumitted (NOV-1999) to the EMBL/GenBank/DDBJ databases. EMBL; AF205284; AAP23405.1;
                                                                                                                                                                                                                                                                                                                                                                                          Query Match
4.5%; Score 138.5; DB 5; Length 782;
Best Local Similarity 21.1%; Pred. No. 0.051;
Matches 110; Conservative 71; Mismatches 176; Indels 165; Gaps
                                                                                                                                                                                                                                                           76 VDDGNINLTDTSTSNK-----SKKGHGRSRVRSASAAAILEEDDSKDDMEF---- 122
                                                                                                                                                                                                                                                                                                     231 PKGKVTILKGIPIKKTKKGCRKSCSGFVQSDSKRESVCNKADAESEPVAQKSQLDRTVCI 290
                                                                                                                                                                                                                                                                                                                                                                                                                                                                          201 TLTPLEELYPTNVNLFNYKYSLNNMEENINILKNEGDLVAQKEEFEYDENMEKAKQDKKK 260
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          --EKI------FQEDTIPRTQIERRKTSLYFSSKYNKEALSPPRRKAFKKWTPPRS 443
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             308 -DEVIEKT---EVSKKTFSGIGFNLTDKEAKVLGVGVTYQEYPETMLYNCPNNSNLFDTI 363
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              ------EVARTADWRDVSELLKPLGLYDLRAK-TIVKFSDEYLTKQWKYPIEL 534
                                                                                                                                                                  123 SLANYLHKNGET----SLKPEDFDF-----TVLSKRGIKSRYKDCSMAALTSHLQNQS- 171
                                                                                                                                                                                                                                    172 -NNSNWNLRTRSKCKKDVFMPPSSSSELQESRGLSNFTSTHLLLKEDEGVDDVNFRKVRK 230
                                                                                                                                                                                                                                                                                                                                   -KASPSVVKTSTPSGTQTSGLKSSS---PSSTKSSSPSN------V 158
                                                                                                                                                                                                                                                                                                                                                                        291 SDAGACGETLSVTSEENSLVKKKERSLSSGSNFCSEQKTSGIINKFCSAKDSEHNEKYED 350
                                                                                                                                    Gaps
                                                                                                                                                                                         18 NVADGINVNGDNNYGKTIINNDFNFDDYNYMTPINKKEFLNSYED--EFSSESFLENKSS 75
                                                                                                                                                                                                                                                                                                                                                                                                                                        TFLESEEI-GTKVEVVERKEHLHT----DILKRGS------EMDNNCSPTRKDFTG
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         261 ALEKIGKOSDEEPFMFSENKFLENQVKERNVAGSFSRFFSK--LNP----FKK----
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               44 PFNLVQETLEHDPWKLLIATIFLNRTSGKMAIPVLWKFLEKYPSA------
                                                                                                                    Best Local Similarity 21.1%; Pred. No. 0.035;
Matches 110; Conservative 71; Mismatches 176; Indels 165;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    Eukaryota; Alveolata; Apicomplexa; Haemosporida; Plasmodium
NCBI_TaxID=5833;
                                                                                                    Length
Li X.R., Luo S.H., Yu X.B., Shan Z.X., Ma C.L.;
Submitted (NOV-1999) to the EMBL/GenBank/DDBJ databases.
EMBL, AF206631; AAF15365.1; -
SEQUENCE 782 AA; 90082 WW; BELF4CF2883903FD CRC64;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            364 ESLOGRIIDIKKRESMISTTFEQOKECLKNMGVLDLELNDTQCKFGT---
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            782 AA; 90041 MW; 27F2EA9BC930434E CRC64;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    01-MAY-2000 (TrEMBLrel. 13, Last sequence update)
01-MAY-2000 (TrEMBLrel. 13, Last annotation update)
RHOPTRY-ASSOCIATED PROTEIN 1.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                535 HGIGKYGNDSYRIF-CVNEWKQVHP-----EDHKLNKYH 567
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                411 -CIGSFGEHHLRLYEFENDLFKFHPNIDYLTLADGYKLQKNH 451
                                                                                                 1 4.6%; Score 140.5; DB 5; Similarity 21.1%; Pred. No. 0.036;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    01-MAY-2000 (TrEMBLrel. 13, Created)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     PRT;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     PRELIMINARY;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         Plasmodium falciparum.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           SEQUENCE FROM N.A.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            STRAIN-FVO;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              Saul A.;
                                                                                                    Query Match
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     090429
                                                                                                                                                                                                                                                                                                                                                                                                                                            351
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              a
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   ð
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 셤
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      q
                                                                                                                                                                                                                                                                                                                                                                                                        q
                                                                                                                                                                                                                                                                                                                                                                                                                                            ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                          염
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           δ
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                ò
     S R R
                                                                                                                                                                                                                                       ŏ
                                                                                                                                                                                                                                                                       g
                                                                                                                                                                                                                                                                                                     ò
                                                                                                                                                                                                                                                                                                                                                                        ò
```

```
172 -NNSNWNLRTRSKCKKDVFMPPSSSSELQESRGLSNFTSTHLLLKEDEGVDDVNFR---- 226
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    76 VDDGNINLTDTSASNK-----SSRKGHGRSRVRSASAAAILEEDDSKDDIEFKASPS 127
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          123 SLANYLHKNGET----SLKPEDFDF-----TVLSKRGIKSRYKDCSMAALTSHLQNQS- 171
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              --KVRKPKGKVTILKGIPIKKTKKGCRKSCSGFVQSDSKRESVCNKADAESEPVAQKSQL 284
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       128 VVKTSTPTGTHT----SSSPSSTKSSS 154
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              201 TLTPLEELYPTNVNLFNYKYSLNNMEENINILKNEGDLVAQKEEFEYDENMEKAKQDKKK 260
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               --EKI--------FOEDTIPRTQIERRKTSLYFSSKYNKEALSPPRRKAFKKWTPPRS 443
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           444 PFNLVQETLFHDPWKLLIATIFLNRTSGKMAIPVLWKFLEKYPSA-------- 488
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     308 -DEVIEKT---EVSKKTFSGIGFNLTDKEAKVLGVGATYQEYPETMLYNCPNNSNLFDTI 363
                                                                                                                                               172 -NNSNWNLRTRSKCKKDVFMPPSSSSELQESRGLSNFTSTHLLLLKEDEGVDDVNFRKVRK 230
                                                                                                                                                                                                                     76 VDDGNINLTDTSTSNK-----SSKKGHGRSRVRSASAAAILEEDDSKDDMEF---- 122
                                                                                                                                                                                                                                                                                               231 PKGKVTILKGIPIKKTKKGCRKSCSGFVQSDSKRESVCNKADAESEPVAQKSQLDRTVCI 290
                                                                                                                                                                                                                                                                                                                                                              123 -KASPSVVKTSTPSGTQTSGLKSSS---PSSTKSSSPSN------V 158
                                                                                                                                                                                                                                                                                                                                                                                                                                               291 SDAGACGETLSVTSEENSLVKKKERSLSSGSNFCSEQKTSGIINKFCSAKDSEHNEKYED 350
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           351 TFLESEEI-GTKVEVVERKEHLHT----DILKRGS-----EMDNNCSPTRKDFTG 395
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               489 ------EVARTADWRDVSELLKPLGLYDLRAK-TIVKFSDEYLTKQWKYPIEL 534
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     18 NYADGINVNGDNNYGKTIIHNDENFDDYNYWTPINKKEFLNSYED--EFSSESFLENKSS 75
123 SLANYLHKNGET----SLKPEDFDF-----TVLSKRGIKSRYKDCSMAALTSHLQNQS- 171
                                                      18 NVADGINVNGDNNYGKTIINNDFNFDDYNYWTPINKKEFLNSYED--EFSSESFLENKSS 75
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     261 ALEKIGKESDEEPFMFSENKFLENQVKERNVAGSFSRFFSK--LNP----FKK----
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    171; Indels 177;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  4.5%; Score 137.5; DB 5; Length 782; 21.8%; Pred. No. 0.06;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                Eukaryota; Alveolata; Apicomplexa; Haemosporida; Plasmodium.
NCBI_TaxID=5854;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             364 ESLQGRIIDIKKRESMISTTFEQQKECLKNMGVLDLELNDTQCKFGT----
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    Shi Y., Lal A.A.;
Submitted (FEB-1995) to the EMBL/GenBank/DDBJ databases
EMBL; U20986; AA63682.1; -.
SEQUENCE 782 AA6, 90160 MW; DE6D1BE2FAC308A9 CRC64;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      01-NOV-1996 (TrEMBLrel. 01, Created)
01-NOV-1996 (TrEMBLrel. 01, Last sequence update)
01-JUN-2000 (TrEMBLrel. 14, Last annotation update)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       535 HGIGKYGNDSYRIF-CVNEWKQVHP------EDHKLNKYH 567
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  65; Mismatches
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        PRT;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       RHOPTRY ASSOCIATED PROTEIN 1.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 Best Local Similarity 21.8 Matches 115; Conservative
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        PRELIMINARY;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     Plasmodium reichenowi
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     SEQUENCE FROM N.A.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         δy
```

27;

Qy	285	DRTVCISDAGACGETLSVTSEENSLVKKKERSLSSGSNFCSEQKTSGIINKFCSAKDSEH 344
Qy	345	NEKYEDTFLESEEI-GTKVEVVERKEHLHTDILKRGSEMDUNCSPT 389 ::
Qy Db	390 1	RKDFTGEKIFQEDTIPRTQIERRKTSLYFSSKYNKEALSPPRRKAFKK 437 :
Oy	438 1	WIPPRSPFNLVQETLFHDPWKLLIATIFLNRTSGRMAIPVLWKFLEKYP 486
QQ	308	.:: : : : : : : : : :
Οy	87	SAEVARTADWRDVSELLKPLGLYDLRAK-TIVKFSDEYLTKQW 528
Op	28	HLFDTISSLQGRLIDIKKRENMISTTFEQQKECLKNMGVLDLELKDTECKFGT 410
Qy	529 1	KYPIELHGIGKYGNDSYRIF-CVNEWKOVHPEDHKLNKYH 567
	1	4 COOL CONTRACTOR OF THE WIND BY THE OF THE
RESULT 0900G6 ID 09	ULT 15 0G6 Q9U0G6	6 PRELIMINARY; PRT; 2209 AA.
A TO TO	01-MAY- 01-MAY- 01-MAY- 01-MAR-	Jo; AY-2000 (TrEMBLrel. 13, Created) AY-2000 (TrEMBLrel. 13, Last sequence update) AR-2001 (TrEMBLrel. 16, Last annotation undate)
DE	VAR, N	14P2.56.
00 00 00 00 00	Plasmc Eukary NCBI_1	Plasmodium falciparum (isolate 3D7). Eukaryota; Alveolata; Apicomplexa; Haemosporida; Plasmodium. NCBL_TaxID=36329;
R P C	[1] SEQUENCE F STRAIN=3D7	SEQUENCE FROM N.A.
RA RA	Devlin Quail	Bowman S., Church Barrell B.;
SQ SR	Submitted EMBL; AL03 SEQUENCE	(FEB-1999 5475; CAB 2209 AA;
Ĉ	erv Mat	
Mae	Best Local	al Similarity 18.8%; Pred. No. 0.14; all Similarity 18.8%; Pred. No. 0.14; all Similarity 18.8%; Mismatches 226; Indels 263; Gaps 37;
Qy	80	SLGDRGAAPTVTSSERLVPDPPNDLRKEDVAMELERVGEDEEQ
Dp	1176 S	
Q 40	54 M	LLQEPIASAQFGATAGTBCRKSVPCGW- 90
3 8	100	CAGDENIIKIGSGIGEDCKNNLPENPSTFKDLEYPTCAKYCRFYKKWI 12
B 5		KKTEYEI
QY	134 T	YELKPEDFDFTVLSKRGIKSRYKDCSMAALTSHLQNQSNNSNWNLRTRSKCKKD 187
pp	1346 G	GTIKFNGGQTFQHTEDCKSCS
Qy	80	VPMPPSSSSELQESRGLSNFTSTHLLLKEDEGVDDVNFRRVR 229
qq		SDNGATGFKGDDLKEACEGKGIFEG
Οy	230 K	KPKGKVTILKGIPIKKTKKGCRKSCSGFVQSDSKRF 265

1498	292	1556	316	1616	375	1662	425	1722	485	1766	542	1817		
: : : :	SVCNKADAESEPVAQKSQLDRTVCISD	9 SHCTKTDQGSTCQNKCQNKCKCVGEWIPKKREEWQQIKDRFLKQYKNDKLDEDFNLRS	AGACGETLSVTSEENS	1557 CLETFLVQIGAAYGEDKFKKVIKLSVFDQSCGCSAIASSQKRNGEYKDAIECMLKKLEEK	7 LSSGSNFCSEQKTSG-IINKFCSAKDSEHNEKYEDTFLESEEIGTKVEVVERKEHLHTDI	7ANKCKEDHSSGEQTEKECQESPSVEDEDDTLHEETEVKAPEICKDV	LKRGSEMDNN - CSP TRKDF TGEKIFQEDTIPRTQIERRKTSLYFSSKYNKE	IKAPTEPEEKGACDPAPTTPKETSPATDSGKETNTEPVTPQDQSPDTKTPKEKGPKPPKS	ALSPPRRKAFKKWTPPRSPFNLVQETLFHDPWKLLIATIFLNRTSGKMAIPVLWKFLEKY	1723 LPQPPRQKREFTPSDWWKVMSASAF-PWTVGVAFMALSYWVMKKK	PSAEVARTADWRDVSEL-LKPLGLYDLRAKT1VK	1767 SKPRVDLESVMEIPQNDYGMPTLKSKNRYVPYSSGKYRGKTYLK-VEGDSGT) DS 544	1818 DS 1819
1439	266	1499	293	1557	317	1617	376	1663	426	1723	486	1767	543	1818
qq	Qy	qq	Ωy	qo	Qy	qq	Qy	qq	QY	qq	Qy	QQ	Qy	qO

Search completed: January 15, 2002, 14:06:31 Job time: 970 sec

		* * * * * * * * * * * * * * * * * * * *
		٠

```
GenCore version 4.5
Copyright (c) 1993 - 2000 Compugen Ltd.
```

OM protein - protein search, using sw model

January 15, 2002, 13:52:31; Run on:

Search time 74.94 Seconds (without alignments) 322.222 Million cell updates/sec

US-09-724-296-37 Perfect score: Title:

1 MIFREGEVSNAMSLWDASPA.......ELSSIRGVKRIGGGALQWKS 317 Sequence:

BLOSUM62 Scoring table:

Gapop 10.0 , Gapext 0.5

219241 Total number of hits satisfying chosen parameters:

219241 seqs, 76174552 residues

Searched:

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0% Maximum Match 100% Listing first 45 summaries

Database :

pir1:* pir2:* pir3:* pir4:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

		de			SUMMARIES	
Result No.	Score	Query	Length	DB	ID	Description
	1650.5		320	7	S55418	od esee Compopue
7	484.5		322	7	397	ָ מוני
m	359.5	21	599	~	S71134	r,
4	345			~	S55262	UV-endonuclease -
S	293.5	17		7	LO.	probable IIV damage
9	104.5	9		7	ത	
7	101	9		~	S74063	
æ	100	9.0		7	\sim	
6			331	7	S54263	
10	96.5		696	7	T38478	ma j
11		5.7	775	7	T22200	hypothetical profe
12	4	٠.	553	~	T21233	
13	94.5		1145	7	A59251	
14	3		802	~	T32377	hypothetical profe
15	3		1036	7	F71565	-
16			674	7	T22733	hypothetical profe
17	92.5		601	7	T02633	
18	92	5.5	428	7	G75064	
19	92		921	~	C81153	
20	91.5		237	7	B82644	5-amino-6-(5-phosp
21	91.5		522	7	G02522	sorting nexin 1 -
22	91.5		1409	-	OFFFCP	copia polyprotein
23	91.5		2334	~	S32920	cell wall-associat
24	91		558	7	T48150	stress-induced pro
25	91		852	~	B84001	
26	91		1020	7	E86165	hypothetical prote
27	91		1284	-	WMVZAI	(r)
28	90.2	5.4	233	Н	R5HG1T	otein
58	90.5	5.4	279	7	T15313	hypothetical prote

hypothetical prote	cytochrome P450 4D	hypothetical profe	Outer membrane pro	retrovirus-related	ATP-dependent seri	myosin-3 heavy cha	hypothetical prote	apolipoprofein A-T	hypothetical prote	bacitracin synthet	hypothetical profe	aspartatetRNA li	type I site-specif	mycosubtilin synth	Ig light chain-bin
T15589	S41192	T32739	E64610	FOHYIH	C82712	T38774	T16816	S29565	T24639	T31677	877357	164203	S02168	T44807	A42808
7	٦	7	7	_	7	7	7	7	7	~	2	7	7	7	7
448	496	625	629	572	848	2104	279	429	553	5255	489	550	1033	5369	719
5.4	5.4	5.4	5.4	5.4	5.4	5.4	5.4	5.4	5.4	5.4	5.4	5.4	5.4	5.4	5.3
											_	_	_		
90.5	90.5	90.5	90.5	06	06	06	89.5	89.5	89.5	89.5	58	58	88	89	88.5

1	
Ē,	α
\supset	ಶ
S	S
Ш	r.
ĸ	U,

UV-endonuclease homolog ywjD - Bacillus subtilis

Cispecies: Bacillus subtilis Cibate: 15-Jul-1995 #sequence_revision 01-Sep-1995 #text_change 15-Oct-1999 Ciracession: S55418: #870060 Richaser, P.: Danchin, A. Submitted to the EMBL Data Library, May 1995 A:Description: Cloning and sequencing of the Bacillus subtilis chromosomal region fro A:Reference number: S55414 A:Reference number: S55414 A:Status: preliminary

A: Molecule type: DNA

A; Residues: 1-320 <GLA>

A;Cross-references: EMBL:249782; NID:9853752; PIDN:CAA89865.1; PID:9853757
R;Kunst, F.; Ogasawara, N.; Moszer, I.; Albertini, A.M.; Alloni, G.; Azevedo, V.; Ber C.; Bron, S.; Broullet, S.; Bruschi, C.V.; Caldwell, B.; Capuano, V.; Carter, N.M.; A.; Ehrlich, S.D.; Emmerson, P.T.; Entian, K.D.; Errington, J.; Fabret, C.; Ferrari, Nature 390, 249-256, 1997
A. Authors: Foulger, D.; Fritz, C.; Fujita, M.; Fujita, Y.; Funa, S.; Galizzi, A.; Gal A. Authors: Eoulger, D.; Krogh, S.; Kunano, M.; Kurita, K.; Lapidus, A.; Hullo, M. Koetter, P.; Koningstein, G.; Krogh, S.; Kunano, M.; Kurita, K.; Lapidus, A.; Lardino A; Authors: Lauber, J.; Lazarevic, V.; Lee, S.M.; Levine, A.; Liu, H.; Masuda, S.; Mau A; Authors: Lauber, J.; Lazarevic, V.; Lee, S.M.; Levine, A.; Liu, H.; Masuda, S.; Mau S, M.; Ogawa, K.; Ogiwara, A.; Odega, B.; Park, S.H.; Parro, V. ; Pohl, T. T.M.; Portete Rieger, M.; Rivolta, C.; Rocha, E.; Roche, B.; Rose, M.; Sadaie, Y.; Sato, T.; Scanl A; Authors: Schleich, S.; Schroeter, R.; Scoffone, F.; Sekiguchi, J.; Sekowska, A.; Se Akuthors: Schleich, A.; Tanaka, T.; Terpstra, P.; Tognoni, A.; Tanaka, A.; Tanaka, T.; Terpstra, P.; Tognoni, A.; Tosato, V.; Uchiya A; Authors: Yoshikawa, H.F.; Zumstein, E.; Yoshikawa, H.; Danchin, A.; Tosato, W.; Voshida A; Title: The complete genome sequence of the Gram-positive bacterium Bacillus subtili A; Accession: B70060

A; Status: nucleic acid sequence not shown; translation not shown DNA A; Molecule type:

A; Residues: 1-320 <KUN>

A;Cross-references: GB:Z99123; GB:AL009126; NID:g2636240; PIDN:CAB15748.1; PID:e11862 A;Experimental source: strain 168

C; Genetics:

99.3%; Score 1650.5; DB 2; Length 320; 99.1%; Pred. No. 1.2e-128; ive 0; Mismatches 0; Indels 3; Ouery Match
Best Local Similarity 99.13
Matches 317; Conservative

Gaps 3;

Ĥ

1 MIFREGEVSNAMSLWDASPAKTLTFARYSKLSKTERKEALLTVTKANLRNTMRTLHYIIG 60 ŏλ Dβ

61 HGIPLYRFSSSIVPLATHPDVMWDFVTPFQKEFREIGELVKTHQLRTSFHPNQFTLFTSP 120 ò Q

~

```
A;Title: A eukaryotic gene encoding an endonuclease that specifically repairs DNA dam A;Reference number: $55262; MUID:95292980
A;Accession: $55262
R;Takao, M.; Yonemasu, R.; Yamamoto, K.; Yasui, A.
Nucleic Acids Res. 24, 1267-1271, 1996
A;Tille: Characterization of a UV endonuclease gene from the fisson yeast Schizosacch A;Reference number: S71134: MUID:96188860
A;Accession: S71134
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            A;Cross-references: EMBL:AL023859; PIDN:CAA19577.1; GSPDB:GN00067; SPDB:SPBC19C7.09c
A;Experimental source: strain 972h-; cosmid c19C7
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  R.Yajima, H.; Takao, M.; Yasuhira, S.; Zhao, J.H.; Ishii, C.; Inoue, H.; Yasui, A.
EMBO J. 14, 2393-2399, 1995
                                                                                                                                                                       A. Status: preliminary; nucleic acid sequence not shown
A. Status: preliminary; nucleic acid sequence not shown
A. Molecule type: mRNA
A. Residues: 1-599 - 4784
A. Cross-references: EMBL: D78571; NID: 91399000; PIDN: BAAl1415.1; PID: 91399001
A. Experimental source: strain SP972
A. Experimental source: strain SP972
A. Experimental source: strain SP972
A. Reference number: 21881
A. Reference number: 21881
A. A. Rocession: 139815
A. Status: preliminary; translated from GB/EMBL/DDBJ
A. Molecule type: DNA
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    C.Species: Neurospora crassa
C.Date: 10-Oct-1995 #sequence_revision 03-Nov-1995 #text_change 20-Jun-2000
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              148 --SVINTHIGGAYGNKDTATAQFHQNIKQLPQEIKERMTLENDDKTYTTEETLQVCEQED 205
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 394 KDAVLIIHLGGTFEGKKETLDRFRKNYQRLSDSVKARLVLENDDVSWSVQDLLPLCQELN 453
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       206 VPFVFDFHHFYANP----DDHADLNVALPRMIKTWERIGLQPKVHLSSPKSEQAI---- 256
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          454 IPLVLDWHHHNIVPGTLREGSLDLMPLIPTIRETWTRKGITQKOHYSESADPTAISGMKR 513
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          46 ANLRNTMRTLHYIIGHGIPLYRFSSSIVPLATHPDVMWDFVTPFQKE-FREIGELVKTHQ 104
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    276 TIQRDGLESVKQLGTQNVLDLIKLVEWNHNFGIHFMRVSSDLFPFASH--AKYGYTLEFA 333
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       90 OKEFREIGELVKTHOLRTSFHPNOFTLFTSPKESVTKNAVTDMAYHYRMLEAMGIADR-- 147
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    26; Gaps
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    Gaps
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          34 TERKEALLTVTKANLRNTM---RTLHYIIGHGIPLYRFSSSIVPLATHPDVMWDFVTPF- 89
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    25;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          21.6%; Score 359.5; DB 2; Length 599; 29.7%; Pred. No. 1e-21; ive 60; Mismatches 111; Indels 25;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              Length 656;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    Indels
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          A; Cross-references: EMBL: D11392; NID: 91526560; PID: 9927215
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              20.8%; Score 345; DB 2; Le 32.3%; Pred. No. 1.8e-20; iive 51; Mismatches 105;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    257 RSHADYVDANFLLERFRQWGTNIDFMIEAKQKDKALLRL 295
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          514 RAHSDRV-----FDFPPCDPTMDLMIEAKEKEQAVFEL 546
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 A;Status: nucleic acid sequence not shown
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     UV-endonuclease - Neurospora crassa
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            Conservative
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        83; Conservative
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 A; Gene: UVDE; SPBC19C7.09c
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     Best Local Similarity
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             Best Local Similarity
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            A; Molecule type: mRNA
A; Residues: 1-656 <YAJ>
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 A; Residues: 1-599 <LYN>
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           C; Accession: S55262
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    A; Map position: 2
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            87;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             Query Match
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     Query Match
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         C; Genetics:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                Matches
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                Matches
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               RESULT
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      S55262
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          ó
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            qq
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      δλ
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     QΩ
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   òλ
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     qq
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 QQ
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               δλ
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             hypothetical protein BH2597 [imported] - Bacillus halodurans (strain C-125)
C;Species: Bacillus halodurans
C;Accession: E83974
R;Takami, H; Nakasone, K; Takaki, Y; Maeno, G; Sasaki, R; Masui, N; Fuji, F; Hira
Nucleic Acids Res. 28, 4317-4331, 2000
A;Title: Complete genome sequence of the alkaliphilic bacterium Bacillus halodurans and
A;Reference number: A83650; MUID:20263314
A;Accession: E83374
A;Accession: E83074
A;Accession: DNA
A;Accession: DNA
A;Accession: C;Graphilianary
A;Accession: C;Graphilianary
A;Cross-references: GB:AP001516; GB:BA000004; NID:g10175192; PIDN:BAB06316.1; GSPDB:GNOG
A;Cross-references: strain C-125
C;Genetics:
A;Gene: BH2597
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               9
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             UV-endonuclease - fission yeast (Schizosaccharomyces pombe)
C;Species: Schizosaccharomyces pombe
C;Date: 21-Apr.1997 #sequence_revision 09-May-1997 #text_change 20-Jun-2000
C;Accession: S71134; T39815
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         231 RMIKTWERIGLQPKVHLSSPKSEQAIRSHADYVDANFLLERFRQWGT----NIDFMIEAK 286
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                59 NLAYQISFFRLSSKLVPLVNHPLTEGWKYELAIAEELQAVGEFASEHQMRIDFHPDHFVV 118
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      117 FTSPKESVTKNAVTDMAYHYRMLEAMGIADRSVINIHIGGAYGNKDTATAQFHQNIKQLP 176
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      177 QEIKERMTLENDDKTYTTEETLQVCEQEDVPFVFDFHHFYANPDDH------ADLNVALP 230
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         179 KSLLSMIMLENDDKSYTIDBVLYLGEKLAIPVVLDIHH------HDVLHRSKSLQETWQ 231
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     58 IIGHGIPLYRFSSSIVPLATHP-DVMWDFVTPFQKEFREIGELVKTHQLRTSFHPNQFTL 116
                                                                                         181 ERMTLENDDKTYTTEETLQVCEQEDVPFVFDFHHFYANPDDHADLNVALPRMIKTWERIG 240
                                                          121 KESVTKNAVTDMAYHYRMLEAMGIADRSVINIHIGGAYGNKDTATAQFHQNIKQLPQEIK 180
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   23; Gaps
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    1 MIFREGEVSNAMSLWDASPAKTLTFARYSKLSKTE---RKEALLTVTKANLRNTMRTLHY 57
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     29.2%; Score 484.5; DB 2; Length 322; 32.2%; Pred. No. 1.9e-32; Live 67; Mismatches 133; Indels 23.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           287 OKDKALLRLMDELSSIRGVKRIGGGALQW 315
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 292 KKDLALFOLMKDLAEYDEITVVSKSAVEF 320
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   ELSSIRGVKRIGGGALQWKS 317
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              301 ELSSIRGVKRIGGGALQWKS 320
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       Matches 106; Conservative
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        Best Local Similarity
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        Query Match
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          RESULT 3
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              ŏ
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             Pp
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           ŏ
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       q
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          ò
                                                                                                                                                                                                                                                                                 g
                                                                                                                                                                                                                                                                                                                                                           ò
                                                                                                                                                                                                                                                                                                                                                                                                                             q
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       δŏ
                                                                                                                                    g
                                                                                                                                                                                                              ò
```

7;

8

Db 277 ANARDIVKMICWNEKYGIRFIRISSEMPPPASHP-VHGYKLAPPASEVLAEAGRVAAELG 335	C;Accession: B59093 R:Okinaka, R.T.; Cloud, K.; Hampton, O.; Hoffmaster, A.R.; Hill, K.K.; Keim, P.; Koeh
OF INCUSSIONAL PROPERTY AND	كاسعداد عادد
QY TOS ENTSFERPRETETTSKENTYTOMATHINGLEAMGLADKSTRINGLEGARGN 100 1 1 1 1 1 1 1 1 1 1	מכונס ליומפשים
161 KDTATAQFHQNIKQLPQEIKERMTLENDDKTYTTEFTLQVCEQEDVPFVFDFHHFYANPD 22	A,Status: preliminary A,Molecule type: DNA A;Residues: 1-315 <0K1>
DB 596 KAATLERFEKRIVARLSQSCKNRLVLENDDVGWTVHDLLFVCEELNIPWLDYHHHNICFD 455 QY 221 DHADLNVALPRMIKTWERIGLQPKVHLSSPKSEQAIRSHADYVDANF 267	A) Cass references: UsikNova44; NID:94094210; FIDN:AMD22322.1; FID:94094234 A.Experimental source: strain Steme A;Note: similar to integrase-recombinase protein (311 aa), MTH893; Methanobacterium t
Db 456 PAHLREGTLDISDPKLQERIANTWKRKGIKQKMHYSEP-CDGAVTPRHRRKHRPRV 510	C;Genetics: A;Genetics: A;Genetics:
Qy 268 ELERFRQWGTNIDFMIEAKQKDKALLRLM 296	A GREGING PARTIE
Db 511 WTLPPCPPDMDLMIEAKDKEQAVFELM 537	Query Match 6.3%; Score 104.5; DB 2; Length 315; Best Local Similarity 22.8%; Pred. No. 0.42; Matches 65; Conservative 41; Mismatches 114; Indels 65; Gaps 10;
RESULT 5 C75350	Qy 32 SKTERKEALLTVTKANLRNTMRTLHYIIGHGIPLYRFSSSIVPLATHPDVMMDFVTPFQK 91
probable UV damage endonuclease - Deinococcus radiodurans (strain K1) C;Species: Deinococcus radiodurans O:Date: 03-hec-1000 #cennance revision 03-hec-1000 #text chance 31-Mar-2000	DD 67 SELEHKRKYSOKYRHRIWVVLKKLHTFLGIVSPLDGIHLSLIPDOSLSDSDFITELEE 124
SOD .	Qy 92 EFREIGELVKTHQLRTSFHPNQFTLFTSPKESVTKNAVTDMAYHYRMLEAMGIADRSVIN 151
M.; Shen, M.; Vamathevan, J.J.; Lam, P.; McDonald, L.; Utterback, T.; Zalewski, C.; Md. S.; Smith, H.O.; Venter, J.C.; Fraser, C.M.	Db 125TLLKQTVLSTKGLTERQAKYRPLIMDRNACI 155
Affile: Genome sequence of the radioresistant bacterium Deinococcus radiodurans RI.	Oy 152 IHIGGAYGNKDTATAQFHQNIKQLPQEIKERMTLENDDKTYTTEE 196
A, reference number: A/JZJO, MUID:ZVOJOOJO A, ACCESSION (77535)	107 PFORDYDE - VEDERH FY ANDROHADI MYALDBAITYTHEFIS OF THE STATE OF THE S
A:Status: preliminary A:Molecule type: DNA A:Desidues: 1-306 /WHT.	1LQVCCQCDDVFF VEDFIN
A.Cross references: GB.RE002022; GB.RE000513; NID:96459590; PIDN:AAF11370.1; PID:9645959 A.Experimental source: strain R1	250 PKSEQAIRSHADYVDANFILERFROWGTNIDFWIEAKQKDKALLR 294
C;Genetics: A;Gene: DRIB A;Map position: 1	264
	RESULT 7
Query Match Query Match Best Local Similarity 30.1%; Pred. No. 1.1e-16; Matches 88: Conservative 51: Mismatches 120: Indels 33: Gaps 11:	S74063 hypothetical protein c0132 - Sulfolobus solfataricus C:Species: Sulfolobus solfataricus
	C; Date: 09-Oct-1997 #sequence_revision 24-Oct-1997 #text_change 20-Jun-2000
QY 21 KTLTFARKSKLSKTERKEALLTVTKANLRNTMRTLHYIGHGIPLYRFSSSIVPLATHHU 80 1:1-1:11 1 1 1 1 1 1 1 Db 46 RTVTLSRYRALSPERREAKLLDLYSSNIKTLRGAADICAHDIRLYRLSSSLPPML D 102	lard, G.; Chan, C.C.Y.; Li
	A,Title: Organizational characteristics and information content of an archaeal genome A,Reference number: S73076; MUID:97055432
1	A; Accession: \$74063 A; Status: nucleic acid sequence not shown; translation not shown
136 YRMLEAMGIADRSVINIHIGGAYGNKDTATAQFHQNIKQLPQEIKERMTLENDDKTYTTE 1	A; Mesidues: 1.2Pe: DAN A; Mesidues: 1.2Ps: SEN> A; Cross-references: EMBL:Y08256; NID:91707679; PIDN:CAA69434.1; PID:91707762
Db 163 ARVMDGLGLA-RTPWNLLLLHGGKGGRGAELAALIPDLPDPVRLRLGLENDERAYSPA 219	A, Experimental source: strain P2 A; Note: the nucleotide sequence was submitted to the EMBL Data Library, September 199
Qy 196 ETLQVCEQEDVPFVFDFHHFYANPDDHADLNVALPRMIKTWERIGLQPKVHL 247	.hermoautotrophicum endonucleas
Qy 248 SSPKSEQAIRSHADYVDANFLLERFRQWGTNIDFM-IEAKQKDKALLRLM 296	«
Db 275 SNGIEGPQDRRHSHLIADFPSAYADVPWIEVEAKGREEAIAALRLM 320	87; Indel
RESULT 6	Qy 95 EIGELVKTHQLRTSFH-PNQFTLFTSPKESVTKNAVTDMAYHYRMLEAMGIADRSVIN 151
1970-1970 protein pXOI-18 - Bacillus anthracis virulence plasmid pXOI C.Species: Bacillus anthracis	ARELGVRLSVHAPYFINLCSEEKEKIEA
C;Date: 12-Nov-1999 #sequence_revision 12-Nov-1999 #text_change 11-May-2000	QY 152 IHIGGAYGNKDTATAQFHQNIKQLPQEIKERMTLENDDKTYTTEETL 198

 Db 113 IH		A; Description: Genetic analysis of the minimal replicon of plasmid pIP417 and compari
,	+	A. Accession: 554263
VQ 991 VQ : IS 071 dd	QVCECEDVPFVPDFHHFYANPDDHADLNVALFRMIKTWERIGLOPRVHLSSPKSEQAI 256 1::	A.Status: preliminary A.Molecule type: DNA A.Residues: 1-331 <-HAGS
Qy 257 RS	RSHADYVDANFLLERFRQWGTNIDFMIEAKQKDKALJRLMDELSS 301	A;Cross-references: EMBL:X86702; NID:q804960; PIDN:CAA60390.1; PID:g804961
Db 222 YR	YRRGKYVDEHIPIDANAPPFEPLAKELLKRDISTTLICESPELERDALKMKEVLER 277	Query Match 5.9%; Score 98; DB 2; Length 331; Best Local similarity 10.0%; Drod No. 1 6.
Qy 302 I	I 302	bocar Similarity 19.3%, Fied. les 52; Conservative 42; Mis
278 L	278	Qy 63 IPLYRESSIVPLATHEDVMMDEVTPEQKEFKEIGELVKTHQLRTSFHPNQFTLFTSPKE 122 :
G83905 hvpothetical	protein BH2047 (imported) - Barillue halzdurane (ettain 2.105)	OY 123 SVTKNAVTDMAYHYRMLEAM-GIADRSVINIHIGGAYGNKDTATAQFHQNIKQLPQEIKE 181
C; Species: B C; Date: 01-D	Sacillus halodurans Dec-2000 #sequence revision 01-Dec-2000 #text change 31-Dec-2000	Db 166TPL1YSIDQLKEMFCVADKYKRVNDFMKKVVE 197
C; Accession: R; Takami, H.	C,Accession: G83905 R;Takami, H.; Nakasone, K.; Takaki, Y.; Maeno, G.; Sasaki, R.; Masui, N.; Fuji, F.; Hira	182 RMTLENDDKTYTTEETLQVCEQ } : :
Nucleic Acid A;Title: Com A:Reference	Is Res. 28, 4317-4331, 2000 splete genome sequence of the alkaliphilic bacterium Bacillus halodurans and number. A 5550. MITEL 2002-2314	198 AAKKELDEVSPYTFEFTPLKSGRKITAIKFYPVYQPEHRDPDL 24
A) Accession: G83905 A) Status: preliminary	. G83905 - eliminary	OY 230 PRMIKIWEKIGLOPEVBHISSPKSEQAIRSHADYDANFLLERFROWGTNIDEMIEAK 286 1
A; Motecule Lype: DNA A; Residues: 1-823 <sto> A; Cross-references: GB:</sto>	Type: Day STD> 1.823 <sto> erences: GB:AP001514; GB:BA000004; NID:910174613; PIDN:BAB05766.1; GSPDB:GN0C</sto>	
A;Experimental C;Genetics: A;Gene: BH2047	source: strain C-125	Db 286SLLPDILGELAILKGKSR 303
Query Matc Best Local	6.0%; Score 100; DB imilarity 23.1%; Pred. No. 4.1;	RESULT 10 T38478 RhoGAP/LIM domain protein - fission yeast (Schizosaccharomyces pombe) C;Species: Schizosaccharomyces pombe
Match	75; Conservative 44; Mismatches 101; Indels 104; Gaps 18;	C;Date: 03-Dec-1999 #sequence_revision 03-Dec-1999 #text_change 03-Dec-1999 C:Accession: T38478
Oy 30 KL Db 332 KL	30 KLSKTERKEALLTVTKANLRNTMRTLHYIIGHGIPLYRFSSSIVPLATHPDV-MW 83	R:Brown, D.; Churcher, C.M.; Wood, V.; Barrell, B.G.; Rajandream, M.A. submitted to the EMBL Data Library, June 1997 A:Reference number: 221796
Qy 84 DF	KESVTKN	A;Accession: T38478 A;Status: preliminary; translated from GB/EMBL/DDBJ
Db 378 QE	QERLEWLWKVKQAYPYWTLGQKQELFHKLESTFTEWTGEQPRRLGG 423	
Oy 142 MG	MGIADERSVINIHIGGAYGNKDTATAQFHQNIKQLPQEIK 180 ::: EIYADRNILYEECHGDINNIKIGGSIKHVIKODVDKWISICAKHGEDBAKORGALAGRIF 483	A;Cross-references: EMBL:29/210; PIDN:CAB10138.1; GSPDB:GN00066; SPDB:SPAC29A4.11 A;Experimental source: strain 972h-; cosmid c29A4 C;Genetics:
181	ERWILENDDKTYTTETLOVCEOEDVPFVFDFHHFYANDDHADL.NVAL.PRWIFFWEFTG 240	A;Wene: SPDB:SPACZ9A4.11 A;Map position: 1
484		5.8%;
Qy 241 LQI	LQPKVHLSSPKSEQAIRSHADYVDANFILERFRQWGTNIDFMIEAKQKDKA 291	Best Local Similarity 18.5%; Pred. No. 10; Matches 60; Conservative 52; Mismatches 99; Indels 113; Gaps 15;
Db 516 QS:	IKTEIEEAITHEVASNPKSVYHLSLDYQDFDRDMAWLTSPDLMI-AKKDDDS 568	OY 17 ASPAKTLTFARYSKLSKTERKEALLTVTKANLRNTMRTLHYIIGHGIPLYRFSSSIV 73
Qy 292 LLI	LLRLMDELSSIRGVKRIGGGALQW 315	/
699	GW.	QY 74 PLATHPDVMWDFVTPFQKEFREIGELVKTHQLRTSFHPNQFTLFTSPKESVTKNAVT 130
RESULT 9 S54263		OV 131 DMAYHYRMLEAMGTADRSVINTHIGGAYGNK PTATAOFHONT 172
rep A protei C;Species: Ba C:Date: 08-I	n - Bacteroides fragilis acteroides fragilis in1-1905 #earnanca revision 21-101-1005 #eart observed 00-00+-1000	: : : : : : : : : : : : : : : : : : :
C; Accession: R; Haggoud, A submitted to	r,	OY 173 KOLPQEIKERMTLENDDKTYTTEETLOVCEQEDVPFVFDFHHFYANDDHA 223
		40/ IALIANLSERLSSSIFUNSAFIKIEDADIVKSARLEICEKFFSFADVIDDFI

oy a	224 DLNVALPRMIKTWERIGLQPKVHLSSPKSEQAIRSHADYV-DANFLLERFRQ-WGTNIDF 281	
g &	SSIRGV 305	
qq	:: :: : : ::: : : : : : : ::: : : :	
RESUL 17222(hypol C; Spk C; Dal C; Spk C; Spk R; McC R; Mc R; M	RESULT 11 T22200 hypothetical protein F44G3.3 - Caenorhabditis elegans C;Species: Caenorhabditis elegans C;Species: Caenorhabditis elegans C;Species: Caenorhabditis elegans C;Accession: T22200 C;Accession: T22200 C;Accession: T22200 A;Reference number: 219529 A;Accession: T22200 A;Acce	
Aa.	<pre>Query Match 5.7%; Score 95; DB 2; Length 775; Best Local Similarity 21.3%; Pred. No. 9.8; Matches 64; Conservative 46; Mismatches 135; Indels 56; Gaps 12;</pre>	
Qy	28 YSKLSKTERKEALLTVTKANLRNTMRTLHYIIGHGIPLYRFSSSIVPLATHPDVMMDFVT 87 :	
Qy	88 PFQKEFREIGELVKTHQLRTSFHPNOFTLFTSPKESVTKNAVTDMAYHYRMLEAMGI 144 :: :	
Qy Dp	145 ADRSVINIHIGGAYGNKDTATAQFHQNIKQLPQEIKERWTLENDDKTYTTEE- 196 :	
Qy Db	197TLQVCEQEDVPFVFDFHHFYANPDDHADLNVALPRMIKTWERIGLOPKVHLSSPKSEQ 254	
Qy Db	255 AIRSHADYVDANFLLERFRQWGTNIDFMIEAKQKDKALLRLMDELSSIRGVKRIGGGALQ 314 1	
Qy Dp	315 W 315 733 W 733	
RESU T212 hypc C;SF C;Da C;Da	RESULT 12 T21233 hypothetical protein F22B3.6 - Caenorhabditis elegans C;Species: Caenorhabditis elegans C;Date: 15-oct-1999 #sequence_revision 15-oct-1999 #text_change 15-oct-1999 C;Accession: T21233 R;Cottage, A.	_
subn A; Re A; Ac	nitted to the EMBL Data Library, January 1996 sference number: 219394 cession: T21233	
A; St A; Mc	Status: preliminary; translated from GB/EMBL/DDBJ Molecule type: DNA Residues: 1-553 <will></will>	

```
A;Description: Molecular analysis of the cDNA coding for an unconventional myosin fro A;Reference number: A59251
A;Accession: A59251
                                                                                                                                                                                                                       10;
A;Cross-references: EMBL:Z68336; PIDN:CAA92738.1; GSPDB:GN00022; CESP:F22B3.6 A;Experimental source: clone F22B3 C;Genetics: A;Gene; CESP:F22B3.6 A;Gene; CESP:F22B3.6
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         C;Species: Acetabularia cliftonii
C;Date: 19-May-2000 #sequence_revision 19-May-2000 #text_change 08-Sep-2000
C;Accession: A59251
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               404 KIPVQQLKT-----VTPLKVPAAE 436
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                159 GNKDTATAQFHQNI-KQLPQEIKERMTLENDDKTYTTEETLQVCEQEDVPFVFDFHHFYA 217
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 50 NTMRTLHYIIGHGIPLYRFSSSIVPLATHPDVMWDFVTPFQKEF...----REIGELV 100
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           349 NAMRTIGMSTGEODSVFRIVAAILHLGN----ITFMGSDEAEFSGSEAEESAQNCADLL 403
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       101 K--THQLRTSFHPNQFTLFTSPKESVTKNAVTDMAYHYRMLEAMGIADRSVINIHIGGAY 158
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 218 NPDDHADLNVALPRMIKTWERIGLQPKVHLSSPKSEQAIR----SHADYVDANFLLERF 272
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             491 NSFEQLCINLAN-----EKLQQQFNHHVLEGEQQQYIAEGISWSYVDFVDNQDCLDLL 543
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               Gaps
                                                                                                                                                                                                                                                                      60 GHGIPLYRFSSSIVPLATHPDVMWDFVTPFQKEFREIGELV---KTHQLRTSFHPNQFTL 116
                                                                                                                                                                                                                                                                                                   117 FTSPKESVTKNAVTDMAYHYRMLEAMGIADRSVINIHIGGAYGNKDTATAQFHQNIKQLP 176
                                                                                                                                                                                                                                                                                                                                                                                                         339 FSSPSPTTOLDDVDEYIYLGROINAQ---NNLMPEIH-----RRRRAAWAAFN-GIKNTT 389
                                                                                                                                                                                                                                                                                                                                                                                                                                                             177 QEIKERMTLEN------DDKTYTTEE---TLQVCEQEDVPFVFDFHHFYANPDDHADLN 226
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       390 DSITDKKIRANLFDSIVLPALTYGSEAWTFTKALSERVRI-----THASLE 435
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      227 VALPRMIKTWERIGLQPKVHLSSPKSEQAIRSHADYVDANFLLERFRQWGTNIDFMIEAK 286
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                436 RRLVGITLTTOOR---ERDLHREDIRTMSLVRD-----PLNFVKKRKLGWAGHV-----AR 482
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    A, Cross-references: GB:U94398; NID:g2051982; PIDN:AAB53062.1; PID:g2051983
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       5.7%; Score 94.5; DB 2; Length 1145; 20.1%; Pred. No. 19;
                                                                                                                                                                          5.7%; Score 94.5; DB 2; Length 553;
20.8%; Pred. No. 6.5;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               A;Status: preliminary; not compared with conceptual translation A;Molecule type: mRNA A;Residues: 1-1145 <VUG>
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               Indels
                                                                                                                                                                                                                            Indels
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               Mismatches 100;
                                                                                                                                                                                                                            41; Mismatches 123;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        A.Gene: myo2
A.Genetic code: SGC5
C.Superfamily: myosin motor domain homology
F.97-799/Domain: myosin motor domain homology <MMO>
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         287 QKDKALLRLMDELSSIRGVKRIGGGALQW 315
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   483 RKDGRWTTLMTEWRPYGWKRPVGRPPMRW 511
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  47;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     R;Vugrek, O.; Menzel, D.
submitted to GenBank, March 1997
                                                                                                                                                                                                     20.8%;
                                                                                                                                                                     Query Match
Best Local Similarity 20.8%;
Warches 56; Conservative
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           Ouery Match
Best Local Similarity 20.1%
                                                                                                               A; Map position:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       C;Genetics
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               RESULT 13
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     qq
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  q
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              q
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             δ
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         qq
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        δ
                                                                                                                                                                                                                                                                                                                          qq
                                                                                                                                                                                                                                                                                                                                                                                                                     ДQ
                                                                                                                                                                                                                                                                                                                                                                                                                                                                    δ
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 qq
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             δy
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            qq
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              ŏ
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     q
```

st

```
Search completed: January 15, 2002, 13:52:32
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       Job time: 401 sec
                                                                                                                                                                                                                                                   Protection protein K10F12.5 - Caenorhabditis elegans
Cipectaes caenorhabditis elegans
New Mondidmann, P.; Beck, C.
Submitted to the RMED bate Library, September 1997
A. Residues relaminary: translated from GB/EMBL/DBJ
A. Accession: T32377
A. Accession: T32377
A. Accession: T32377
A. Accession: T3237
A
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    117 FTSPKESVTKNAVTDMAYHYRMLEAMGIADRSVINIHIGGAYGNKDTATAQFHQNIKQLP 176
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 591 FGSPSPTTQLDDVDEYIYLGRQINAQ---NNLMPEIH----RRRRAAWAAFN-GIKNTT 641
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                177 QEIKERMTLEN------DDKTYTTEE---TLQVCEQEDVPFVFDFHHFYANPDDHADLN 226
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    227 VALPRMIKTWERIGLQPKVHLSSPKSEQAIRSHADYVDANFLLERFRQWGTNIDFMIEAK 286
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      60 GHGIPLYRFSSSIVPLATHPDVMWDFVTPFQKEFREIGELV---KTHQLRTSFHPNQFTL 116
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              642 DSITDKKIRANLFDSIVLPALTYGSEAWTFTKALSERVRI------THASLE 687
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              688 RRLVGITLTQQR----ERDLHREDIRTMSLVRD-----PLNFVKRKLGWAGHV-----AR 734
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             Gaps
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              Query Match 5.6%; Score 93.5; DB 2; Length 805; Best Local Similarity 20.8%; Pred. No. 14; Matches 56; Conservative 41; Mismatches 123; Indels 49;
                                                         544 EGGG-----KNKAKGIFPLIDEACKMPNV 567
273 ROWGTNIDFMIEAKQKDKALLRLMDELSSIRGV 305
    δ
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             qq
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  ŏ
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       Qγ
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                δŏ
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                Óχ
```

```
A;Title: Genome sequence of an obligate intracellular pathogen of humans: Chlamydia t A;Reference number: A71570; MUID:99000809
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              A;Status: preliminary
A;Molecule type: DNA
A;Residues: 1-1036 <ARN>
A;Cross·references: GB:AE001277; GB:AE001273, NID:g3328408; PIDN:AAC67609.1; PID:g332
                                                                                                                                                                                                                                                                                                                                                     RiStephens, R.S.; Kalman, S.; Lammel, C.J.; Fan, J.; Marathe, R.; Aravind, L.; Mitche Science 282, 754-759, 1998
                                                                                                                                                                                                                        probable isoleucine--tRNA ligase (EC 6.1.1.5) - Chlamydia trachomatis (serotype D,
                                                                                                                                                                                                                                                       C:Species: Chlamydia trachomatis
C:Date: 13-Sep-1998 #sequence_revision 13-Sep-1998 #text_change 24-Nov-1999
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       5.6%; Score 93.5; DB 2; Length 1036;
19.7%; Pred. No. 20;
tive 42; Mismatches 104; Indels 87; Gaps
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 71 SIVPLATHPDVMW-----DFVT-----PFQKFFREIGELVKTHQLRTSFHP 111
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          229 SNMALAVHPELTYVRIKDKESGDEYILGQESLPRWFPDRESYEWIGQLSGKSLVGQSYEP 288
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    165 TAQFH----TTEETL 198
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    332 ACQEHNVPLVCPVDNQGCYTAEVKDFVGEYIKSADKGIARR--LKNENKLFYQGTVRHRY 389
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              199 QVCEQEDVPFVFDFHHFYANPDDHADLNVALPRMIKTWERIGLQPKVHLSSPKSEQAIRS 258
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           112 NQFTLFTSPKESVTKNAVTDMAYHYRMLEAMGIADRSVIN-----IHIGGAYGNKDTA 164
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        390 PFCWRTDSPLIYKAVNSW----FVAVEKVKSKMLKANESIHWTPE-HIKQGRFGKWLEG 443
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            259 HADYVDANFLLERFRQWGTNIDFMIEAKQKDKALLRL--MDELSSIRGVK 306
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  444 ARDWA----ISRNRYWGTPIPIW---RSDDGELLVIGSIQELEALSGOK 485
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  A; Experimental source: serotype D, strain UW-3/Cx
287 QKDKALLRLMDELSSIRGVKRIGGGALQW 315
                                                            735 RKDGRWTTLMTEWRPYGWKRPVGRPPMRW 763
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      C; Superfamily: isoleucine -- tRNA ligase
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    57; Conservative
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    Query Match
Best Local Similarity
Matches 57; Conserva
                                                                                                                                                                                                                                                                                                                          C; Accession: F71565
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        C; Keywords: ligase
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           A; Gene: iles
   δy
                                                         q
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       òγ
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        qq
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              QQ
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  ò
```

·		

NO PA

GenCore version 4.5 Copyright (c) 1993 - 2000 Compugen Ltd.

OM protein - protein search, using sw model

Run on:

January 15, 2002, 14:07:17; Search time 45.07 Seconds (without alignments) 257.882 Million cell updates/sec

US-09-724-296-37 1662 1 MIFREGEVSNAMSLWDASPA.....ELSSIRGVKRIGGGALQWKS(317) Title: Perfect score: Sequence:

BLOSUM62 Gapop 10.0 , Gapext 0.5 Scoring table:

100059 seqs, 36664827 residues Searched:

100059 Total number of hits satisfying chosen parameters:

Minimum DB seq length: 0 Maximum DB seq length: 200000000

Post-processing: Minimum Match 0% Maximum Match 100% Listing first 45 summaries

SwissProt_39:* Database :

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Description	20 00 11 10 00 PM PAGE 10 10 10 10 10 10 10 10 10 10 10 10 10	373 bacillus			013596 homo sanies							014157 schizosacch	_		P47282 mycoplasma		_				~				9 gallus ga		Q62431 mus musculu	_	P77994 thermotoga	~	3344	7a0	39 schizosacc
Ω	YWJD BACSU		SYICHLTR	C4D2 DROME	SNX1 HUMAN	COPI_DROME	WAPA BACSU	ATI_COWPX	RL1_THEMA	GAG_IPHA	MYH2_HUMAN	MYS3_SCHPO	APA4_MACFA	BACA_BACLI	SYD_MYCGE	T1R1_ECOLI	HEMO_OPSTA	VPH1_NEUCR	KCCD_RAT	PUR2_YEAST	LIN1_HUMAN	RHO_THEMA	FOJO_DROME	UL52_VZVD	MYPF_CHICK	HEMZ_PASMU	DRII_MOUSE	MATB_NEUCR	RPSD_THEMA	CPV1_RAT	MASY_EMENI	SYT_CHLPN	CUT7_SCHPO
DB	:		Н																						Η,								П
Length	320	964	1036	501	522	1409	2334	1284	233	572	1941	2104	429	5255	220	1033	582	826	533	805	1259	427	583	1083	1131	326	100	382	399	208	528	635	1085
% Query Match	99.		5.6							5.4	5.4	5.4	5.4	5.4	5.4	5.4	ω.		2.5	5.2	5.2	5.2	2.5	5.2	5.2	ν. Τ.	 	 	7.5.	5.1	 	5.1	5.1
Score	5.		93.5					O١	90.5	06	90	۰۰,	89.5	œ. ,	68	≖,	88.5	ή,	٠.	ė	٠,	98	80	9 0	a c	0 0	0	•	٠		84.5	٠	84.5
Result No.	٦	7	m	4	2	9	7	∞ ·	σ,	01,	Ι;	77	T .	14	15	10	17	80° 6	19	707	77	77	23	4 7	070	0 7 0	7 0 0	9 0	אַנּ	300	31	3.2	33

57 saccharomyc		n homo sapion		72 saccharomyc	59 vaccinia vi	50 saccharomic	20 saccinatollyc	43 hadiopus tae	43 Dactiles su	of thermonlasm	80 saccharomyc
P404	014647	9/60	1000	P11972	1 P C d	0500		1000	2551	7227	P32380
YIO9_YEAST	CHD2 HUMAN	COXZ HUMAN	METX YEAST	SST2 YEAST	ATII VACCV	YMS1 YEAST	EP45 XENTA	GUTE BACSU	RBP2 PLAVB	METE THEAC	NUF1_YEAST
_	П	Н	\leftarrow	Н	۲	Η	_	-	H	Н	\vdash
1679	1739	276	649	698	724	843	436	829	1251	343	944
5.1	5.1	5.1	5.1	5.1	5.1	5.1	5.0	5.0	5.0	5.0	5.0
84.5	84.5	84	84	84	84	84	83.5	83.5	83.5	83	83
34	32	36	37	38	39	40	4.1	42	43	44	45

ALIGNMENTS

"WAD_BACSU STANDARD; PRT; 320 AA. "WAD_BACSU STANDARD; PRT; 320 AA. 101-NOV-1995 (Rel. 32, Last sequence update) 102-AUG-2001 (Rel. 40, Last amnotation update) 103-AUG-2001 (Rel. 40, Last amnotation update) 104-AUG-2001 (Rel. 40, Last amnotation update) 105-AUG-2001 (Rel. 40, Last amnotation update) 106-AUG-2001 (Rel. 40, Last amnotation update) 107-AUG-2001 (Rel. 40, Last amnotation update) 108-CHILLS STAID-1423; 108-CHILLS STAID-1423; 109-CHILLS STAID-1		ERMTLENDDKTYTTEETLQVCEQEDVPFVFDFHHFVANPDDHADLNVALPRMIKTWERIG 240 [1]
PACSU YWJD_BACSU STANDARD; P45864, 01-NOV-1995 (Rel. 32, Last 20-AuG-2001 (Rel. 36, Last HYPOTHETICAL 36.9 KDA PROTY WJD. BACillus Subtilis. Bacillus/Staphylococcus 91 Bacillus/Staphylococcus 91 NCBI_TaxID=1423, [1] SEQUENCE FROM N.A. STRAIN=168; Glaser P., de la Fuente V. Submitted (MAY-1995) to th This SWISS-PROT entry is c between the Swiss Institute the European Bioinformatic use by non-profit institute between the Swiss Institute the European Bioinformatic use by non-profit institute or send and this statemen entities requires a licens or send an email to licens cor send an email to licens or send an email to licens from this Swiss Institute Swiss Institute The Buchelical protein; Comp SEQUENCE 320 AA; 36896 try Match SEQUENCE 320 AA; 36896 try Match IMTERCEVSNAMSIMDASPAK ililililililililililililililililililil	HGIPLYRESSSIVPL HGIPLYRFSSSIVPL KESVTKNAVTDMAYH HIHHHHHHHHHHHHHHHHHHHHHHHHHHHHHHHHHH	ERMTLENDDKTYTTE
RESULT 1 YWJD_BACSU TD YWJD_BACSU AC P45864; DT 01-NOV-1 DT 01-NOV	7 7	Oy 181 Db 181 Oy 241

How olid they kno so alleto it?

```
Davis R.W.;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    SEQUENCE
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           BINDING
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   Matches
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         q
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           Dp
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              ÓΫ
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               qq
  SON THE TEXT TO BE THE TO SON THE TEXT TO SON 
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             Ŏλ
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       Qγ
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 This SWISS-PROT entry is copyright. It is produced through a collaboration
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See http://www.isb-sib.ch/announce/or send an email to license@isb-sib.ch).
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         EMBL outstation
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               between the Swiss Institute of Bloinformatics and the EMBL outstation
the European Bloinformatics Institute. There are no restrictions on
use by non-profit institutions as long as its content is in no
                                                                                                                                                                                                                                                                                       15-JUL-1998 (Rel. 36, Last sequence update)
15-JUL-1998 (Rel. 36, Last annotation update)
15-JUL-1998 (Rel. 36, Last annotation update)
17-JUL-1998 (Rel. 36, Last annotation update)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      1 Similarity 22.1%; Pred. No. 1.2;
55, Conservative 44; Mismatches 84; Indels 66; Gaps
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            238 RIGLQPKVHLSSP-----KSEQAIRS------HADYVDANFLLERFRQWGTNIDFMIEA 285
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               152 -----IHIGGAYGNKDTATAQFHQNIKQLPQ-EIKERMTLENDDKTY-----TTEET- 197
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               612 DAGKQFINIAGVNTSPDGAANTVWHMSKRTPEYNILHLINLVNNDQNWRNSGNQPTAQTN 671
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          198 ----LQVCEQEDVPFVFDFHHFYANPDDHADLNVALP-------RMIKTWE 237
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           672 LATKVYIGAEETITGV----YAASPDHNQGATQSLPFTTGTDSSGSYISFTVPSLEYWS 726
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  95 EIGELVKTHQLRTSFHPNQFTLFTSPKESVTKNAVTDMAYHYRMLEAMG--IADRSVIN- 151
241 LQPKVHLSSPKSEQAIRSHADYVDANFLLPLLERFROWGTNIDFMIEAKQKDKALLRLMD 300
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  Oguma T., Kurokawa T., Tobe K., Kitao S., Kobayashi M.;
Submitted (NOV-1996) to the EMBL/GenBank/DDBJ databases.
-!- FUNCTION: PRODUCES CYCLOISOMALTOOLIGOSACCHARIDE FROM DEXTRAN.
-!- SIMILARITY: BELONGS TO FAMILY 66 OF GLYCOSYL HYDROLASES.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             6.2%; Score 103; DB 1; Length 964; 22.1%; Pred. No. 1.2;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                CYCLOISOMALTOOLIGOSACCHARIDE
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    964 AA; 107208 MW; 8849CDC5E2DE9A68 CRC64;
                                                                                                                                                                                                                                                                                                                                                                                                                                                               Bacteria; Firmicutes; Bacillus/Clostridium group;
Bacillus/Staphylococcus group; Bacillus.
NCBL_TaxID=1397;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         GLUCANOTRANSFERASE.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   1036 AA.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  Transferase; Glycosyltransferase; Signal.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             POTENTIAL.
                                                                                                                                                                                                                                         PRT;
                                                           298 ELSSIRGVKRIGGGALQWKS 317
                                                                                         301 ELSSIRGVKRIGGGALQWKS 320
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    EMBL; D88360; BAA13595.1; -
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   STANDARD;
                                                                                                                                                                                                                                              STANDARD;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              Best Local Similarity
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        286 KOKDKALLR 294
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    784 SSKDDYVLR 792
                                                                                                                                                                                                                                                                                                                                                                                                                                                  Bacillus circulans.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         SEQUENCE FROM N.A.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 STRAIN=U-155
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           SYI_CHLTR
                                                                                                                                                                                                                                              CTA2_BACCI
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            SEQUENCE
                                                                                                                                                                                                                                                                                                                                                                                                   CITASE).
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     Query Match
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      Matches
                                                                                                                                                                                        RESULT 2
CTA2_BACCI
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           SYI_CHLTR
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     RESULT
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         Π
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     qq
                                                                                                                                                                                                                                              δλ
```

```
13;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See http://www.isb-sib.ch/announce/or send an email to license@isb-sib.ch).
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         Stephens R.S., Kalman S., Lammel C.J., Fan J., Marathe R., Aravind L., Mitchell W.P., Olinger L., Tatusov R.L., Zhao Q., Koonin E.V.,
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   199 QVCEQEDVPFVFDFHHFYANPDDHADLNVALPRMIKTWERIGLQPKVHLSSPKSEQAIRS 258
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         390 PFCWRTDSPLIYKAVNSW----FVAVEKVKSKMLKANESIHWTPE-HIKOGRFGKWLEG 443
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            PRINTS; PR00984; TRNASYNTHILE.
PROSITE; PS00178; AA_TRNA_LIGASE_I; 1.
Aminoacyl-tRNA synthetase; Protein biosynthesis; Ligase; ATP-binding;
Aminoacyl-tRNA synthetase; Protein biosynthesis; Ligase; ATP-binding;
Amtal-binding; Zinc; Complete proteome.
SITE 46 56 "HIGH" REGION.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                165 TAQEH----TTEETL 198
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      332 ACCEHNUPLVCPVDNOGCYTAEVKDFVGEYIKSADKGIARR--LKNENKLFYOGTVRHRY 389
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       71 SIVPLATHPDVMW------DFVT-----PFQKEFREIGELVKTHQLRTSFHP 111
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          112 NOFTLETSPKESVTKNAVTDMAYHYRMLEAMGIADRSVIN-----IHIGGAYGNKDTA 164
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   289 -LFPYFQDKKE--------LEAFRILPADFIEESEGTGIVHMAPAFGEADFF 331
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                229 SNMALAVHPELTYVRIKDKESGDEYILGQESLPRWFPDRESYEWIGQLSGKSLVGQSYEP 288
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    Chlamydia trachomatis.";
Science 287.754-759(1998).
-!- CATALYTIC ACTIVITY: ATP + L-ISOLEUCINE + TRNA(ILE) = AMP +
PYROPHOSPHATE + L-ISOLEUCYL-TRNA(ILE).
-!- COPACTOR: BINDS ONE ZINC ION (BY SIMILARITY).
-!- COPACTOR: BINDS ONE ZINC ION (BY SIMILARITY).
-!- SUBUNIT: MONOMER (BY SIMILARITY).
-!- SUBCELLOLAR LOCATION: CYTOPLASMIC.
-!- SIMILARITY: BELONGS TO CLASS: I AMINOACYL-TRNA SYNTHETASE FAMILY.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               "Genome sequence of an obligate intracellular pathogen of humans:
20-AUG-2001 (Rel. 40, Created)
20-AUG-2001 (Rel. 40, Last sequence update)
20-AUG-2001 (Rel. 40, Last annotation update)
30-AUG-2001 (Rel. 40, Last annotation update)
ISOLEUCYL-TRNA SYNTHETASE (EC 6.1.1.5) (ISOLEUCINE--TRNA LIGASE)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        5.6%; Score 93.5; DB 1; Length 1036; 19.7%; Pred. No. 8;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              259 HADYVDANFLLERFRQWGTNIDFMIEAKQKDKALLRL--MDELSSIRGVK 306
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    "KMSKS" REGION.
ATP (BY SIMILARITY).
MW; DC4420FD37343EFD CRC64;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             42; Mismátches 104;
                                                                                                                                                                                                                                    Chlamydia trachomatis.
Bacteria; Chlamydiales; Chlamydiaceae; Chlamydia.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 InterPro; IPR002300; tRNA-synt_la.
InterPro; IPR001412; tRNA-synt_I.
InterPro; IPR002301; tRNA-synt_ile.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                    MEDLINE=99000809; PubMed=9784136;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 592 592 ATI
1036 AA; 118985 MW;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            EMBL; AE001277; AAC67609.1; -.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            Pfam; PF00133; tRNA-synt_1;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         57; Conservative
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                593
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                Query Match
Best Local Similarity
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             HSSP: P56690; 11LE
                                                                                                                                                                                                                                                                                                                                                                                           SEQUENCE FROM N.A.
                                                                                                                                                                                                                                                                                                                                                                                                                                    STRAIN=D/UW-3/CX;
                                                                                                                                                                                                                                                                                                                 NCBI_TaxID=813;
                                                                                                                                                                                                   ILES OR CT019.
```

RAGADELE-20196006; PubMed=10731132;
RAGADELE-20196006; PubMed=10731132;
RA Adams M.D., Celniker S.E., Holt R.A., Evans C.A., Gocayne J.D.,
RA Admanatides P.G., Scherer S.E., Li P.W., Hoskins R.A., Galle R.F.,
RA George R.A., Richards S. Ashburner M., Henderson S.N.,
RA George R.A., Richards S. Ashburner M., Henderson S.N.,
RA George R.A., Barder B.G., Helt G., Champe M., Pfeiffer B.D.,
RA Bardon R.C., Rogers Y.-H.C., Blazej R.G., Champe M., Pfeiffer B.D.,
RA Abril J.F., Agbayani A., An H.-J., Andrews-Pfannkoch C., Baldwin D.,
RA Ballew R.M., Basu A., Baxendale J., Bayraktaroglu L., Beasley E.M.,
RA Ballew R.M., Basu A., Baxendale J., Brokstein P., Brottier P.,
RA Burtis R.C., Busam D.A., Butler H., Cadieu E., Center A., Chandra I.,
RA Borkova D., Botchan M.R., Bouck J., Brokstein P., Brottier P.,
RA Borkova D., Botchan M.R., Buller H., Cadieu E., Center A., Cawley S., Dahlke C., Davenport L.B., Davies P.,
RA Gooson K., Doup L.E., Downes M., Dugan-Rocha S., Dunkov B.C., Dunn P.,
RA Durbin K.J. Evangelista C.C., Ferraz C., Ferriera S., Flerischmann W.,
RA Gooson K., Gong F., Gorrell J.H., Gu Z., Guan P., Harris M.,
RA Harris N.L., Harvey D., Helman T.J., Hernandez J.R., Houck J.,
RA Harris N.L., Marvey D., Helman T.J., Hernisch J.A., Ketchum K.A.,
Jalali M., Kalush F., Karpen G.H., Ke Z., Kennison J.A., Ketchum R.A.
Jasko P., Lei Y., Levitsky A.A., Li J., Li Z., Liang Y., Lin X.,
Alankilov G., Hilshina N.V., Mobarry C., Morris J., Mosheria A.,
Merkulov G., Hilshina N.V., Mobarry C., Murphy D.M., Mosher D.L.,
RA Palazzolo M., Pittman G.S., Panl S., Pollard J., Puri V., Reese M.G., C402_DROME STANDARD; PRT; 501 AA.
C402_DROME STANDARD; Created)
15-DEC-1998 (Rel. 37, Created)
20-AUG-2001 (Rel. 40, Last sequence update)
20-AUG-2001 (Rel. 40, Last sequence update)
20-AUG-2001 (Rel. 40, Last annotation update)
CYPOLENDME P650 4D2 (EC 1.14---) (CYPIVD2).
CYPOLENDME P650 4D2 (EC 1.14---) (CYPIVD2).
CYPOLENDME Melanogaster (Fruit fly).
Prerygota; Metazoa; Arthropoda; Tracheata; Hexapoda; Insecta;
Pterygota; Mooplera; Endopterygota; Diptera; Brachycera; Muscomorpha;
Ephydroidea; Drosophilidae; Drosophila. Benos P.V. Gatt M.K., Ashburner M., Murphy L., Harris D.,
Barrell B.G., Ferraz C., Vidal S., Brun C., Demailles J., Cadieu E.,
Dreano S., Gloux S., Lelaure V., Mortier S., Galibert F., Borkova D.,
Manana B., Kafatos F.C., Louis C., Siden-Kiamos I., Bolshakov S.,
Papagiannakis G., Spanos L., Cox S., Madueno E., de Pablos B.,
Modolell J., Peter A., Schoettler P., Werner M., Mourkloti F.,
Beinert N., Dowe G., Schaefer U., Jaeckle H., Bucheton A.,
Callister D.M., Campbell L.A., Darlamitsou A., Henderson N.S.,
McMillan P.J., Salles C., Tait E.A., Valenti P., Saunders R.D.C., Frolov M.V., Alatortsev V.E.; "Cluster of cytochrome P450 genes on the X chromosome of Drosophila "From sequence to chromosome: the tip of the X chromosome of D. 444 ARDWA----ISRNRYWGTPIPIW---RSDDGELLVIGSIQELEALSGQK 485 MEDLINE=20196011; Pubmed=10731137; MEDLINE=94296569; PubMed=8024706; DNA Cell Biol. 13:663-668(1994). Science 287:2220-2222(2000). SEQUENCE FROM N.A. SEQUENCE FROM N.A. SEQUENCE FROM N.A. STRAIN=OREGON-R; STRAIN-BERKELEY; STRAIN=OREGON-R melanogaster. melanogaster C4D2_DROME RESULT a

```
This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation the Buropean Bloinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities required.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See http://www.isb-sib.ch/announce/
Reinert K., Remington K., Saunders R.D.C., Scheeler F., Shen H., Shue B.C., Siden-Kiamos I., Simpson M., Skupski M.P., Smith T., Spier E., Spradling A.C., Stapleton M., Strong R., Sun E., Svirskas R., Tector C., Turner R., Venter E., Wang A.H., Wang X., Wang Z.-Y., Wassarman D.A., Weinstock G.M., Weissenbach J., Walliams S.M., Woodage T., Worley K.C., Wu D., Yang S., Yao Q.A., Ye J., Yeh R.-F., Zaveri J.S., Zhan M., Zhang G., Zhao Q., Zheng L., Zheng X.H., Zhong F.N., Zhong W., Zhou X., Zhu S., Zhu X., Smith H.O., Gibbs R.A., Myers E.W., Rubin G.M., Venter J.C.;
The genome sequence of Drosophila melanogaster.",
                                                                                                                                                                                                                                                                                       "Evidence for non-neutral evolution around the cytochrome p450 gene cluster on the Drosophila melanogaster X chromosome.";
Submitted (OCT-1997) to the EMBL/GenBank/DDBJ databases.
-!- FUNCTION: INVOLVED IN THE METABOLISM OF INSECT HORMONES AND IN THE BREAKDOWN OF SYNTHETIC INSECTICIBES (BY SIMILARITY).
-!- CATALYTIC ACTIVITY: RH + REDUCED FLAVOPROTEIN + O(2) = ROH + OXIDIZED FLAVOPROTEIN + H(2)0.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       LRSANGVHLGLKPRA -> CGRPTAFILA (IN REF.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        HEME (BY SIMILARITY).
K -> M (IN STRAINS CAM-44, CAM-48 AND
BERKELEY).
                                                                                                                                                                                                                                                                                                                                                                                                                           -!- SUBCELLULAR LOCATION: MEMBRANE-BOUND. ENDOPLASMIC RETICULUM
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                MISSING (IN REF. 2).
I -> A (IN REF. 1; CAA80549).
A -> R (IN REF. 1).
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                PROSITE; PS00086; CYTOCHROME_P450; 1.
Oxidoreductase; Monooxygenase; Membrane; Heme; Microsome;
Endoplasmic reticulum; Polymorphism.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                 -!- SIMILARITY: BELONGS TO THE CYTOCHROME P450 FAMILY.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             8F8D98DBF39FD1A2 CRC64;
                                                                                                                                                                                                                                                     STRAIN=VARIOUS STRAINS;
Phillips K.S., Begun D.J., Aquadro C.F.;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    or send an email to license@isb-sib.ch).
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        EMBL, AL009194; CAR15691; EMBL, AL009194; CAR15691; EMBL, AF017006; AAR711691; EMBL, AF017009; AAR711701; EMBL, AF017009; AAR711711; EMBL, AF017011; AAR711741; EMBL, AF017011; AAR711741; EMBL, AF017012; AAR711761; EMBL, AF017014; AAR711771; EMBL, AF017015; AAR711771; EMBL, AF017014; AAR711771; EMBL, AF017015; AAR711781; EMBL, AF017016; AAR711781; EMBL, AF017016; AAR711781; EMBL, AF017016; AAR711791;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          EMBL; AF017017; AAB71180.1; -. EMBL; AF017018; AAB71181.1; -. FlyBase; FBgn0011576; Cyp4d2. InterPro; IPR001128; Cyt_P450.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           57914 MW;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          EMBL; X75955; CAA53568.1; -.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              CAA80549.1; -.
                                                                                                                                                                                                                                   SEQUENCE OF 4-501 FROM N.A.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              PR00464; EP450II.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      Pfam; PF00067; p450; 1.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        449
163
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            PRINTS; PR00385; P450
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           501 AA;
                                                                                                                                                                                                                                                                                                                                                                                                                                             (POTENTIAL).
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                30
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   160
487
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              223005
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           SEQUENCE
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                CONFLICT
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   CONFLICT
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       CONFLICT
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              PRINTS;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        BINDING
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            VARIANT
```

12;

91; Indels 89; Gaps

5.6%; Score 92.5; DB 1; Length 501; 19.8%; Pred. No. 3.6;

43; Mismatches

Best Local Similarity 19.89 Matches 55; Conservative

Query Match

4

```
This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the FMBL outstation the Buropean Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        entities requires a license agreement (See http://www.isb-sib.ch/announce/
or send an email to license@isb-sib.ch).
                                                                                                                                                                                                                                                   225 LNV-----ALPRMIKT--WERIGLOPKVHLSSPKSEQAIRSHADYVDA----NFLLE 270
60 GHGI-----PLYRFSSSIVPLATHPDVMWDFVTPFQKEFREIGELVKTHQLRT-SFHPNQ 113
                                                                                                                          87 IEFVLSSQQHITKNNL-----YKLL------NCWLGDGLLMSTGRKWHGRRII 129
                                                                                                                                                                                                           166 INIFPVICLTALDIIAETAMGTKINAQKNPNL-------PYVQAVNDVTNILIK 212
                                                                                 114 FTLFTSPKESVTKNAVTDMAYHYRMLEAMGIADRSVINIHIGGA-----YGNKDTA 164
                                                                                                                                                                    165 TAQFHQNIKQLPQEIKERMTLENDDKTYTTEETLQVCEQEDVPFVFDFHHFYANPDDHAD 224
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               MEDLINE=99038232; PubMed=9819414;
Haft C.R., de la Luz Sierra M., Barr V.A., Haft D.H., Taylor S.I.;
Haft C.R., de la Luz Sierra M., Barr V.A., Haft D.H., Taylor S.I.;
"Identification of a family of sorting nexin molecules and
characterization of their association with receptors.";
Mol. Cell. Biol. 18:7278-7287(1998).
-!- FUNCTION: MAY BE INVOLVED IN SEVERAL STAGES OF INTRACELLULAR
TRAFFICKING. PLAYS A ROLE IN TARGETING LIGAND-ACTIVATED EGER TO
THE LYSOSOMES FOR DEGRADATION AFTER ENDOCYTOSIS FROM THE CELL
SURFACE AND RELEASE FROM THE GOLGI.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         MEDLINE-96208617; PubMed-8638121;
Kurten R.C., Cadena D.L., Gill G.N.;
"Enhanced degradation of EGF receptors by a sorting nexin, SNX1.";
Science 272:1008-1010(1996).
                          27 GNGILPGPRPLPFEGNLLMYRGLDPEQIMDFVKKNQRKYGRLYRVWILHQLAVFSTDPRD
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   Homo sapiens (Human).
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     ISOFORMS; 1 (SHOWN HERE) AND 1A; ARE
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      Transport; Protein transport; Golgi stack; Alternative splicing.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                213 RFIHAWQRVDWIFRLTQPTEAKRQDKAIKVMHDFTENI 250
                                                                                                                                                                                                                                                                                                                                      271 RFRQWGTNIDFMI ---- EAKQKDKALLRLMDELSSI 302
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  SNXI_HUMAN STANDARD; PRT; 522 AA. 013596; 060750; 060751; 20-AUG-2001 (Rel. 40, Created) 20-AUG-2001 (Rel. 40, Last sequence update) 20-AUG-2001 (Rel. 40, Last annotation update)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              SEQUENCE FROM N.A., AND ALTERNATIVE SPLICING.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  AF065483; AAC17182.1; -. AF065484; AAC17183.1; -.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              EMBL; U53225; AAA98672.1; -.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          -!- ALTERNATIVE PRODUCTS: 2
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           InterPro; IPR001683; PX. Pfam; PF00787; PX; 1. SMART; SM00312; PX; 1.
                                                                                                                                                                                                                130 TPTFHFKI------
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        SEQUENCE FROM N.A.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              NCBI_TaxID=9606;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         SORTING NEXIN 1.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                   SNX1_HUMAN
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           a
                                                                                                                                                                                                                                                                                                                                      ŏ
                                         g
                                                                                       ò
                                                                                                                            9
                                                                                                                                                                      ρý
                                                                                                                                                                                                                                                       Ω
                                                                                                                                                                                                                                                                                               Q
```

```
16;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           MEDLINE-85240569; PubMed-2409449;
Emori Y., Shiba T., Kanaya S., Inouye S., Yuki S., Saigo K.;
"The nucleotide sequences of copia and copia-related RNA in Drosophila
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  Eukaryota; Metazoa; Arthropoda; Tracheata; Hexapoda; Insecta;
Pterygota; Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;
Ephydroidea; Drosophilidae; Drosophila.
                                                                                                                                                                                                                                                                     77 THPDVMWDEVTPFQKEFREIGELVKTHQLRTSFHPNQFTLFTSPKESVTKNAV----TDM 132
                                                                                                                                                                                                                                                                                                                                                                                                                  186 ENDDKTYTTEETLQVCEQEDVPFVFDFHHFYANPD-----DH----ADLNVALPRM 232
                                                                                                                                                                                                                                                                                                                                                                                                                                                   356 SSEDNTALSRALSQLAEVEEK--IEQLHQEQANNDFFLLAELLSDYIRLLAIVRAAFDQR 413
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         233 IKTWERIGLQPKVHLSSPKSEQAIRSHADYVDANFLLERFRQ-----WGTNIDFMIEAK 286
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             414 MKTWQR---------WQDAQATLQKKREAEARLLWANKPDKLQQAK 450
                                                                                                                                                                                                                                   194 YEKLSEKHSQNGFIVPPSPEKSLIGMTKVKVGKEDSSSAEFLEKRRAALERYLQRIV--- 250
                                                                                                                                                                                                                                                                                                       251 NHPTMLQD----PDVREFLEKEELPRAVGTQTLSGAGLLKMFNKATDAVSKMTIKMNESDI 307
                                                                                                                                                                                                                                                                                                                                            133 AYHYRMLEAMGIADR----SVINIHIGGAYGNKDTA--TAQFHQNIKQLPQEIKERMTL 185
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    Mount S.M., Rubin G.M.; Roomles of the Drosophila transposable element copia; homology between copia and retroviral proteins."; Mol. Cell. Biol. 5:1630-1638(1985).
                                                                                                                                                              Gaps
                                                                                                                                                                                              28 YSKLSKTER------KEALLTVTKANL-RNTMRTLHYIIGHGIPLYRFSSSIVPLA 76
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     Miller K., Resenbaum J., Zbrzezna V., Pogo A.O.; "The nucleotide sequence of Drosophila melanogaster copia-specific 2.1-kb mRNA.";
                                                                                                                                                                                                                                                                                                                                                                                308 WFEEKLOEVECEEORLRKLHAVVETLVN---HRKELALNTAOFAKSLAMLG-----
                                                                                                                                                            106; Indels 107;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            SEQUENCE FROM N.A., SEQUENCE OF 2-10, AND ALTERNATIVE SPLICING
                                                                                                                            DB 1; Length 522;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       P04146; 024585; 024586; 024280; 024555; 024587; 01-NOV-1986 (Rel. 03, Created) 20-A0G-2001 (Rel. 40, Last sequence update) 20-A0G-2001 (Rel. 40, Last annotation update) copia PROTEIN (CONTAINS: COPIA VLP PROTEIN; COPIA PROTEASE
                                  S -> P (IN REF. 1).
S -> P (IN REF. 1).
FC8BB93977A5AE70 CRC64;
                   MISSING (IN ISOFORM 1A).
                                                                                                                                           ; Pred. No. 4.6; 54; Mismatches
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               1409 AA
                                                                                                                          5.5%; Score 91.5; 19.3%; Pred. No. 4.6
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   287 QKDKALLRLMDELSSIRGVKRIGGGALQWKS 317
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      451 -------DEI------LEWES 458
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            Nucleic Acids Res. 17:2134-2134(1989).
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         Drosophila melanogaster (Fruit fly).
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          SEQUENCE FROM N.A. (SHORT ISOFORM).
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               PRT;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    SEQUENCE FROM N.A. (LONG ISOFORM).
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           MEDLINE=89183629; PubMed=2538806;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     MEDLINE=85267679; PubMed=2410772;
                                                                        59049 MW;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     virus-like particles.";
Nature 315:773-776(1985).
                                                                                                                                                                64; Conservative
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               STANDARD;
265
155
117
211
145 2
91 1
117 1
211 2
522 AA;
                                                                                                                                                 Similarity
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 NCBI_TaxID=7227;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          (EC 3.4.23.-)].
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   COPI_DROME
                                                                                                                             Query Match
                    VARSPLIC
                                                                          SEQUENCE
                                        CONFLICT
                                                         CONFLICT
                                                                                                                                                  Best Local
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             COPIA.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              COPI_DROME
                                                                                                                                                                  Matches
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    g
                                                                                                                                                                                                                                                                                                                                                                                                                                                          q
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 δλ
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       Qγ
     FT
FT
FT
SQ
                                                                                                                                                                                                                                                                             ò
                                                                                                                                                                                                                                                                                                               qq
                                                                                                                                                                                                                                                                                                                                                 ŏ
                                                                                                                                                                                                                                                                                                                                                                                  qq
                                                                                                                                                                                                                                                                                                                                                                                                                        Q
```

```
This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way
                                                                                                                                                                                                                               modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See http://www.isb-sib.ch/announce/or send an email to license@isb-sib.ch).
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       STTGYLFKMFDFNLICWNTKRONS -> VOOGIYSKCLILI
                                                                                                               -!- ALTERNATIVE PRODUCTS: 2 ISOFORMS; A LONG FORM (SHOWN HERE) AND A SHORT FORM; ARE PRODUCED BY ALTERNATIVE SPLICING.
-!- SIMILARITY: THE PROTEASE BELONGS TO PEPTIDASE FAMILY All.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     SFVGIQRDRTQ (IN VARIANT COPIA-RELATED).
MISSING (IN VARIANT COPIA-RELATED).
D->A: LOSS OF ACTIVITY.
                                            toshioka K., Honma H., Zushi M., Kondo S., Togashi S., Miyake T.,
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               Transposable element; Hydrolase; Aspartyl protease; ATP-binding;
SEQUENCE FROM N.A. (SHORT ISOFORM), AND MUTAGENESIS OF ASP-292.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             N -> K (IN REF. 1).
E -> Q (IN REF. 1).
S -> N (IN REF. 1).
H -> Q (IN REF. 1).
I -> V (IN REF. 1).
C -> E (IN REF. 2; CAA26447).
V -> C (IN REF. 2; CAA26447).
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        Alternative splicing; Polymorphism.

1 270 COPIA VLP PROTEIN (POTENTIAL).

771 1409 COPIA PROTEASE (POTENTIAL).
                                                                      "Virus-like particle formation of Drosophila copia through
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         (IN SHORT ISOFORM).
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   BE89440763A47691 CRC64;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           PROTEASE (BY SIMILARITY)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 PS00141; ASP_PROTEASE; FALSE_NEG.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                        InterPro, IPR001969; Asp_protease.
InterPro, IPR001784; Bunya_nucleocap.
InterPro, IPR001835; Enterotoxin_B.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          MISSIMG
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       Enterotoxin_B.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  InterPro; IPR001225; FA_desaturase...
InterPro; IPR001884; Rve.
InterPro; IPR001878; Znf CCHC.
Pfam; PF00952; Bunya_nucleocap; 1.
                             MEDLINE=90151630; PubMed=1689241;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            PF01376; Enterotoxin_B; 1.
PF00487; FA_desaturase; 1.
PF00665; rve; 1.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     AA; 162817 MW;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      PF00098; zf-CCHC; 1.
5; PR00939; C2HCZNFINGER.
: SM00343; ZnF_C2HC; 1.
                                                                                                                                                                                                                                                                                                                                   X02600; CAA26446.1;

X02600; CAA2647.1;

X13719; CAA31997.1;

X54147; CAA38086.1;

M11240; AAA74498.1;
                                                                                                                                                                                                                                                                                             EMBL; X04456; CAA28054.1; -. EMBL; X02599; CAA26444.1; -.
                                                                                                                                                                                                                                                                                                                          X02599; CAA26445.1; -.
                                                                                       alytic processing.";
9:535-541(1990).
                                                                                                                                                                                                                                                                                                                                                                                                                                                          FlyBase; FBgn0000349; copia.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         292
1374
1288
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  1409
                                                                                                                                                                                                                                                                                                                                                                                                                                S14835.
                                                                                                                                                                                                                                                                                                                                                                                                             PIR; A03324; OFFFCP
                                                                                                                                                                                                                                                                                                                                                                                                                                            MEROPS; A11.001
                                                                                       autocatalytic
                                                                                                                                                                                                                                                                                                                                                                                                                                S14835;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                Polyprotein;
                                                          Shiba T.;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        ACT_SITE
VARSPLIC
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    PRINTS;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                SEQUENCE
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  PROSITE;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              CONFLICT
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               CONFLICT
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            CONFLICT
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            CONFLICT
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         CONFLICT
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        CONFLICT
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      CONFLICT
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       /ARIANT
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  VARIANT
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  MUTAGEN
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     SMART;
                                                                                                                                                                                                                                                                                                                                                       EMBL;
                                                                                                                                                                                                                                                                                                                                                                                  EMBL;
                                                                                                                                                                                                                                                                                                                                                                     EMBL;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            CHAIN
                                                                                                                                                                                                                                                                                                                                                                                                 EMBL;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               Pfam;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         Pfam;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            Pfam;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         Pfam;
```

```
12;
                 Gaps
                                         21 KTLTFARYSKLSKTERKEALLTVTKANLRNTMRTLHYIIGHGIPLYRFSSSIVPLATHPD 80
               71;
                 85; Indels
                                                             <del>..</del>
; Pred. No. 17;
39; Mismatches
                                                          _
=
=
=
            52; Conservative
```

5.5%; Score 91.5; DB 1; Length 1409;

21.1%;

Query Match Best Local Similarity

Matches

δy

```
This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation-the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See http://www.isb-sib.ch/announce/
                                                                                                                                                                                                180 KE-----RMTLENDDKTY-----TTEETLQVCEQEDVPFVFD---FHHFYANP-- 219
                                                                                                                                                                                                                         Yoshida K.-I., Sano H., Seki S., Oda M., Fujimura M., Fujita Y.; "Cloning and sequencing of a 29 kb region of the Bacillus subtilis genome containing the hut and wapa loci."; Microbiology 141:337-343(1995).

-!- FUNCTON: STILL UNKNOWN. NOT INVOLVED IN CELL MEMBRANE METABOLISM, MOTILITY, SECRETION OR DIFFERENTIATION.

-!- SUBCELLULAR LOCATION: MEMBRANE-ASSOCIATED. MAY BE RELEASED
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                SIMILARITY: THE REPEATED MOTIF IN THE C-TERMINUS HAS SOME SIMILARITY TO THE REPEAT IN E.COLI RHS GROUP OF PROTEINS (RHSA-D).
                                                                                                                     125 TKNAVTDMAYHYRMLEAMGIADRSVINIHIGGAYGNKDTATAQFHQNIKQL-----PQEI 179
                                                                                                                                                          698 ARDVVVD---ETNMVNSRAVKFETVF-----LKDSKESENKNFPNDSRKIIOTEFPNES 748
    -----YLINR-IP----SRALVDSSKTPY 637
                                       81 VMWDFVTPFOKEFREIGELVKTH-----V 124
                                                                             638 EMWHNKKPYLKHLRVFGATVYVHIKNKQGKFDDKSFKSIFVGYEPNGFKLWDAVNEKFIV 697
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           HAS THREE
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             INTO THE MEDIUM.
-!- DOMAIN: HAS TWO LIGAND-BINDING DOMAINS; THE N-TERMINUS, HAS THR 101 AA REPEATS WHICH ARE RESPONSIBLE FOR CELL WALL BINDING; THE C-TERMINUS CONSISTS OF TWO BLOCKS OF RESIDUES WITH A CONSERVED
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         Foster S.J.;
Molecular analysis of three major wall-associated proteins of
Bacillus subtilis 168: evidence for processing of the product of
gene encoding a 258 kba precursor two-domain ligand-binding
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   Bacteria; Firmicutes; Bacillus/Clostridium group;
                                                                                                                                                                                                                                                                                                                                                                                                                                                       01-0CT-1994 (Rel. 30, Created)
01-0CT-1994 (Rel. 30, Last sequence update)
20-AUG-2001 (Rel. 40, Last annotation update)
                                                                                                                                                                                                                                                                                                                                                                                                                          PRT; 2334 AA.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      Bacillus/Staphylococcus group; Bacillus.
NCBL_TaxID=1423;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      or send an email to license@isb-sib.ch).
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        WALL-ASSOCIATED PROTEIN PRECURSOR.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               STRAIN=168 / BGSC1A1;
MEDLINE=95219088; PubMed=7704263;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        MEDLINE=93302506; PubMed=8316082;
595 KARTMVSGAKLDKSFWGEAVLTAT---
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         Microbiol. 8:299-310(1993).
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           EMBL; L05634; AAA22883.1; -.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          EMBL; D31856; BAA06656.1; -
EMBL; D29985; BAA06260.1; -
EMBL; D83026; BAA11683.1; -
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   MOTIF REPEATED 31 TIMES.
                                                                                                                                                                                                                                                                                                                                                                                                                        STANDARD;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                Bacillus subtilis.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              SEQUENCE FROM N.A.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               SEQUENCE FROM N.A.
                                                                                                                                                                                                                                                                              220 ----DDH 222
                                                                                                                                                                                                                                                                                                                   809 KRKRDDH 815
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         WAPA OR N17G
                                                                                                                                                                                                                                                                                                                                                                                                                      WAPA_BACSU
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   STRAIN=168;
                                                                                                                                                                                                                                                                                                                                                                                                 WAPA_BACSU
                                     δy
                                                                           q
                                                                                                                   òγ
                                                                                                                                                        q
                                                                                                                                                                                                                                    Db
                                                                                                                                                                                                                                                                                                                   qq
                                                                                                                                                                                                                                                                                                                                                                                                                                      òγ
                                                                                                                                                                                                                                                                              Qγ
```

```
13;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              KTYT--TEETLQVCEQEDVPFVFDFHHFYANPDDHADLNVALPRMIKTWER-----IG 240
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            119 KE--STSKQVETEN------AILNSDEQKQMKNGLYATFEHNDHKVTYSLAEAKGP 166
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           167 NKTSLTPKDTSADYKTDSNEIVYPDVFPNIDLQTFTFNEN-----IKEDLVLHQYNGY 219
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  220 NTFTFQLKTDLQAKEQEDGSIDFS------DEKGKVVFSVPKPFMTDSKLDELSGEVE 271
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            66 KDINKEGEITSER----TENTKLYYEGDGV--YKQEVYLDPIHTKETPDADWEDISPEL 118
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 91 KEFREIGELVKTHQLRTSFHPNQFTLFTSPKESVTKNAV-----TDMAYHYRMLEAMGI 144
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       -----NIHIGGAYGNKDTATAQFHQNIKQLPQEIKERMTLE--NDD 189
                                                                                                                    31 X 21 AA APPROXIMATE TANDEM REPEATS OF
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      67; Gaps
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         33 KTERKEALLTVTKANLRNTMRTLHYIIGHGIPLYRFSSSIVPLATH--PDVMWDFVTPFQ 90
                                              Cell wall; Repeat; Signal; Complete proteome.

OR 32 POTEMNIAL).

CHAIN 29 234 WALL.ASSOCIATED PROTEIN.

DOMAIN 504 869 3 X 101 AA APPROXIMATE TANDEM REPEATS.
                                                                                                                                X(4)-G-X(4)-[YF]-X-D-X(2)-G-X(4).
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    DB 1; Length 2334;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       35; Mismatches 115; Indels
                                                                                                                                                                                                                                                                                                                                                                                                                                                     B75138CCD278BAA3 CRC64;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  RSDKVSYKLEKNEEGYLLHLT-ADENWLKDPERVYPVSID 310
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      241 LQPKVHLSSPKSEQAIRSHADYVDANFLLERFRQWGTNID 280
                                                                                                                                                                                                                                                                                                                                                                       2-24 (APPROXIMATE).
2-25.
2-26.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    5.5%; Score 91.5; 1
22.5%; Pred. No. 34;
                                                                                                                                                                                                                              2-10.
2-11.
2-12.
2-13.
                                                                                                                                                                                                                                                                       2-14.
2-15.
2-16.
2-17.
2-18.
2-19.
2-20.
                                                                                                                                                                      2-4.
2-5.
2-6.
2-8.
                                                                                                                                                                                                                                                                                                                                                                                                                                                      258329 MW;
EMBL; 299124; CAB15959.1; -.
        PIR; S32920; S32920.
Subtilist; BG10797; wapA.
InterPro; IPR003305; CBD_6.
Pfam; PF02018; CBD_6; 1.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       Conservative
                                                                                                                                                                                                                                                                                                     1814
1839
1859
1880
1906
                                                                                                                    2139
                                                                                                                                                                                         1148
1169
1193
                                                                                                                                                                                                                                          1665
                                                                                                                                                                                                                                                    9897
                                                                                                                                                                                                                                                              1709
                                                                                                                                                                                                                                                                        1730
                                                                                                                                                                                                                                                                                                                                                                                                                                                      2334 AA;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                Query Match
Best Local Similarity
Matches 63; Conserva
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         145 ADRSVI-----
                                                                                                                                                                                                                                           646
                                                                                                                                                                                                                                                                                                                                                                                                                                                      SEQUENCE
                                                                                                                                                                                                                                 REPEAT
REPEAT
REPEAT
                                                                                                                                                                                                                                                                                                     REPEAT
REPEAT
                                                                                                                                                                                                                                                                                                                                  REPEAT
REPEAT
                                                                                                                                                                                                                                                                                                                                                                REPEAT
REPEAT
                                                                                                                                                                                                                                                                                                                                                                                            REPEAT
REPEAT
                                                                                                                                                                                                                                                                                                                         REPEAT
                                                                                                                                                                                                                                                                                                                                                       REPEAT
                                                                                                                                                                                                                                                                                                                                                                                    REPEAT
                                                                                                                                                                                                                                                                                                                                                                                                                  REPEAT
                                                                                                                                                                                                                                                                                                                                                                                                                          REPEAT
                                                                                                                                                                                                                                                                                                                                                                                                                                    REPEAT
                                                                                                                                                                                                                                                                                                                                                                                                                                             REPEAT
                                                                                                                      DOMAIN
                                                                                                                                                                                                                        REPEAT
                                                                                                                                                                                                                                                                REPEAT
                                                                                                                                                                                                                                                                        REPEAT
                                                                                                                                                                                                                                                                                   REPEAT
                                                                                                                                                                                                                                                                                            REPEAT
                                                                                         REPEAT
                                                                                                                                                                                  REPEAT
                                                                                                                                                                                          REPEAT
                                                                                                                                                                                                      REPEAT
                                                                                                                                                                                                              REPEAT
                                                                                                   REPEAT
                                                                                                            REPEAT
                                                                                                                                         REPEAT
                                                                                                                                                    REPEAT
                                                                                                                                                             REPEAT
                                                                                                                                                                        REPEAT
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                190
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 q
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      οy
 q
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         ŏ
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           qq
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                δ
```

```
This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss institute of Bioinformatics and the EMBL outstation. The European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See http://www.isb-sib.ch/announce/or send an email to license@isb-sib.ch).
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          15;
                                                                                                                                                                                                                                                                                                                                                                                                                                                          -:- FUNCTION: MATURE VIRUSES ARE OCCLUDED INTO THE ATI, AND IT HAS BEEN ASSUMED THAT SUCH BODIES PROTECT THE VIRUS DURING DISSEMINATION FROM ANIMAL.

-:- MISCELLANEOUS: A CHARACTERISTIC FEATURE OF ATI IS THE FORMATION OF ILARGE MASSES WITH NO SURROUNDING MEMBRANES IN THE CYTOPLASM OF INFECTED CELLS.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       5.5%; Score 91; DB 1; Length 1284;
19.4%; Pred. No. 17;
tive 50; Mismatches 99; Indels 100; Gaps
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        56 HYIIGHGIPLY-----RESSSIVPLATHPDVMWDFVTPFQKEFREIGELVKTHQLRT 107
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       190 HYLLVSGSELFINVAYGPASFSSPI----SVPDYIMEALT----FKACDHIMKSGDLK- 239
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        108 SFHPNQFTLFTSPKESVTKNAVTDMAYHYRMLEAM~--GIADR-----SVINIHIGG 156
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             240 -----YTYAFTKKVKDLFNIKSDSVYQYVRLHEMSYDGVSEDTDDDDEVFAILNLSI-- 291
                                                                                                                                                                                                                                                                             Funahashi S., Sato T., Shida H.; "Cloning and characterization of the gene encoding the major protein
                                                                                                                                                           Viruses; dsDNA viruses, no RNA stage; Poxviridae; Chordopoxvirinae;
                                                                                                                                                                                                                                                                                                                                                                                             Patel D.D., Pickup D.J.;
"Messenger RNAs of a strongly-expressed late gene of cowpox virus contain 5'-terminal poly(A) sequences.";
EMBO J. 6:3787-3794(1987).
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         10 X APPROXIMATE TANDEM REPEATS.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             F7904C9ElDE8D012 CRC64;
                                                                                                                                                                                                                                                                                                           of the A-type inclusion body of cowpox virus.";
J. Gen. Virol. 69:35-47(1988).
                                                    Ploovs;
01-AUG-1990 (Rel. 15, Created)
01-AUG-1990 (Rel. 15, Last sequence update)
01-NOV-1990 (Rel. 16, Last annotation update)
                             PRT; 1284 AA.
                                                                                                                                                                                                                                                                                                                                                             SEQUENCE OF 1-109 FROM N.A. MEDLINE=88111568; Pubmed=2828037;
                                                                                                                                                                                                                                                          MEDLINE=88089536; PubMed=2826668;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             150329 MW;
                                                                                                          A-TYPE INCLUSION PROTEIN (ATI)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            EMBL; D00319; BAA00222.1; -. EMBL; X06343; CAA29650.1; -.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            Local Similarity 19.4 es 60; Conservative
                               STANDARD;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                Repeat
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             1284 AA;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             PIR; JQ0006; WMVZAI
                                                                                                                                             Cowpox virus (CPV)
                                                                                                                                                                                                                              SEQUENCE FROM N.A.
                                                                                                                                                                                            NCBI_TaxID=10243;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             721
752
781
812
812
872
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                Late protein;
                                                                                                                                                                             Orthopoxvirus.
                                                                                                                                                                                                                                                 STRAIN=CPR06
                           ATI_COWPX
P16602;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             SEQUENCE
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              Query Match
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                REPEAT
REPEAT
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             REPEAT
REPEAT
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              DOMAIN
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               REPEAT
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            REPEAT
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                REPEAT
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               REPEAT
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             REPEAT
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              REPEAT
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             Matches
             ATI_COWPX
RESILL'E
                                           qq
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               ÖΣ
```

```
This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation the European Bioinformatics Institute. There are no restrictions on its muse by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See http://www.isb-sib.ch/announce/or send an email to license@isb-sib.ch).
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       TABLE 399:323-329(1999).

-i - FUNCTION: THIS PROTEIN BINDS DIRECTLY TO 23S RIBOSOWAL RNA AND IS LOCATED IN THE BRIGHORHOOD OF THE SITE WHERE ELONGATION FACTOR TU IS BOUND TO THE RIBOSOWE (BY SIMILARITY).

-i - SIMILARITY: BELONGS TO THE LIP FAMILY OF RIBOSOMAL PROTEINS.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      Melson K.B., Clayton R.A., Gill S.R., Gwinn M.L., Dodson R.J.,
Haft D.H., Hickey B.K., Peterson J.D., Nelson W.C., Ketchum K.A.,
McDonald L., Utterback T.R., Malek J.A., Linher K.D., Garrett M.M.,
Stewart A.M., Cotton M.D., Pratt M.S., Phillips C.A., Richardson D.,
Heidelberg J., Sutton G.G., Fleischmann R.D., Eisen J.A., White O.,
Salzberg S.L., Smith H.O., Venter J.C., Fraser C.M.;
"Exidence for lateral gene transfer between Archaea and Bacteria from
genome sequence of Themotoga maritima.";
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        translation component genes in the extremely thermophilic eubacterium
                       292 -----DSSVDRYRNRVLLLTPEVASLRKEYSDVEPDYKYLMDEE------VP-AYDK 336
157 AYGNKDTATAQFHQNIKQLPQEI----KERMTLENDDKTYTTEETLQVCEQEDVPFVFDF 212
                                                                                                                                                             262 YV-----DANFLLERFRQW-----GTNIDFMIEAKQKDKAL----LRLMDE 298
                                                                                                                                                                                    213 H------HFYANPDDHADLNVALPRMIKTWERIGLQPKVHLSSPKSEQAIRSHAD 261
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  Dennis P.P., Liao D.; "The organization and expression of essential transcription
                                                                                                                     337 HLPKPITNTGIEEPHATGGDKEEQEQQPVKVVQSKPDDGITPYNPFEDP
                                                                                                                                                                                                                                                                                                                                                                                                               01-DEC-1992 (Rel. 24, Created)
01-DEC-1992 (Rel. 24, Last sequence update)
20-AUG-2001 (Rel. 40, Last annotation update)
                                                                                                                                                                                                                                                                                                                                                                                233 AA.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           Phermotoga maritima.";
J. Biol. Chem. 267:22787-22797(1992).
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         Thermotogales; Thermotoga.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                InterPro: IPR002143; Ribosomal_L1.
Pfam; PF00687; Ribosomal_L1; 1.
ProDom; PD001314; Ribosomal_L1; 1.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        STRAIN-MSB8 / DSM 3109;
MEDLINE-99287316; PubMed-10360571;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           MEDLINE=93054590; PubMed=1429627;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             EMBL; 211839; CAA77860.1; -.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                EMBL; AE001723; AAD35538.1;
                                                                                                                                                                                                                                                                                                                                                                                STANDARD;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                50S RIBOSOMAL PROTEIN L1
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   FROM N.A.
SB8 / DSM 3109;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   PIR; S19900; R5HG1T.
PIR; C44466; C44466.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         HSSP; P27150; 1AD2.
TIGR; TM0455; ...
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        Thermotoga maritima
                                                                                                                                                                                                                                          299 LSSIRGVKR 307
                                                                                                                                                                                                                                                                   || || ::|
447 LSKIRTLER 455
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       SEQUENCE FROM N.A.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                NCBI_TaxID=2336;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       RPLA OR TM0455.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             STRAIN-MSB8 ,
                                                                                                                                                                                                                                                                                                                                                                              RL1_THEMA
P29393;
                                                                                                                                                                                                                                                                                                                                                             RL1_THEMA
                                                                                                                                                                                                                                                                                                                                                                                                     g
                                                                                                                                                                                                                                            ŏ
                                                                                                                                                                                                                                                                                  g
                                                                                                                     g
```

```
16;
                                                                                                                       .
8
                                                                                                                                                         137 RMLEAMGIADRS-----VINIHIGGAYGNKDTATAQFHQNIKQLPQEIKERMTLENDD 189
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   SEQUENCE FROM N.A.
MEDLINE-65264989; PubMed=2991563;
Ono M., Toh H., Miyata T., Awaya T.;
"Nucleotide sequence of the Syrian hamster intracisternal A-particle
quene: close evolutionary relationship of type A particle gene to
types B and D oncovirus genes.";
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     48; Mismatches 123; Indels 120; Gaps
                                                                                                                                                                                                                           190 KTYTTEETLOVCEQEDVPFVFDFHHFYANPDDHADLNVALPRMIKTWERIGLQPK--VHL 247
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            12 QTVLKQRDLKIAPRTLQNFWKEVDRV----APWYACSGSLTVAS------53
                                                                                                                       59; Gaps
                                                                                                                                                                                                                                                         39 -----DETIELHIQTGID-----YRKPEQHIRGTIVLP-----HGTGKEVKVLVFA 79
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      33 KTERKEALLTVTKANLRNTMRTLHYIIGHGIPLYRFSSSIVPLATHPDVMWDFVTPFQKE 92
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               Retroid viruses; Retroviridae; Intracisternal A-particles.
                                                                                                                                                                                                                                                                                                               248 SSPKSEQAIRSHADYVDANFLLERFRQWGTNIDFMIEAKQKDKALLRLMDELSSIRG 304
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      -!- SIMILARITY: BELONGS TÓ PEPTIDASE FAMILY A2; ALSO KNOWN AS THE RETROPEPSIN FAMILY.
                                                                                      DB 1; Length 233;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                5.4%; Score 90; DB 1; Length 572; 19.4%; Pred. No. 6.9;
                                                                                                                       51; Indels
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          SMART; SMU0343; ZNF_C2HC; 2.
PROSITE; PS00141; ASP_PROTEASE; FALSE_NEG.
PROSITE; PS0175; ASP_PROT_BERTROV; 1.
Core_protein; Polyprotein; Hydrolase; Aspartyl protease.
PROSITE; PS01199; RIBOSOMAL_L1; 1.
Ribosomal protein; rRNA-binding; Complete proteome.
SEQUENCE 233 AA; 25932 MW; F115A70F5C193108 CRC64;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               572 AA; 63930 MW; 3FFD8FD3A1EFD47C CRC64;
                                                                                                                                                                                            7 RYLEARKLVDRTKYYDLDEAIEL-----VKKTATAKF----
                                                                                                                                                                                                                                                                                                                                                                                                                                                       23-OCT-1986 (Rel. 02, Created)
23-OCT-1986 (Rel. 02, Last sequence update)
30-MAY-2000 (Rel. 39, Last annotation update)
RETROVIRUS-RELATED GAG POLYPROTEIN.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               Hamster intracisternal a-particle (IAP-H18)
                                                                                      5.4%; Score 90.5; Dl
23.2%; Pred. No. 1.9;
                                                                                                                       41; Conservative 26; Mismatches
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  POTENTIAL
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              PROTEASE.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              InterPro; IPR001995; Asp_prot_retrov.
InterPro; IPR001969; Asp_protease.
                                                                                                                                                                                                                                                                                                                                                                                                                        PRT;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        PRINTS; PR00939; C2HCZNFINGER. SMART; SM00343; ZnF_C2HC; 2.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 InterPro; IPR003322; Gag_pl0.
InterPro; IPR000721; Gag_p24.
InterPro; IPR001878; Znf_CCHC.
Pfam: PF02337; Gag_pl0; 1.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           55:387-394(1985).
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  Pfam; PF02337; Gag_pl0; 1. Pfam; PF00607; gag_p24; 2.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    Local Similarity 19.49
nes 70; Conservative
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     Pfam; PF00077; rvp; 1.
Pfam; PF00098; zf-CCHC; 1
                                                                                                                                                                                                                                                                                                                                                                                                                        STANDARD;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           PIR; A03952; FOHYIH.
HSSP; P05888; LAAF.
MEROPS; A02.016; -.
                                                                                                        Similarity
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              NCBI_TaxID=11752;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  491
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  ACT_SITE
SEQUENCE
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           J. Virol.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      Query Match
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       Viruses;
                                                                                                                                                                                                                                                                                                                                                                                                                        GAG_IPHA
                                                                                         Query Match
Best Local :
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 CHAIN
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      Matches
                                                                                                                            Matches
                                                                                                                                                                                                                                                                                                                                                                                                       GAG_IPHA
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            q
                                                                                                                                                                                                                                                                                                                                                                                                                                          δλ
                                                                                                                                                                                                                                 ΟŊ
                                                                                                                                                                                                                                                                  Dp
                                                                                                                                                                                                                                                                                                   δλ
     SOM
```

```
TISSUE-skeletal muscle;
MEDLINE=99318869; PubMed=10388558;
Weiss A., Schiaffino S., Leinwand L.A.;
"Comparative sequence analysis of the complete human sarcomeric myosin heavy chain family: implications for functional diversity.";
J. Mol. Biol. 290:61-75(1999).
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       258 SHADYVDANFLLERFRQWGTNIDF-MIEAKQKDKALLRLMDELSSIRGVKRIGGGALQWK 316
                            --KLGKELDRKHA-----EGDLCLGTKAIWKLVKNCLEDEACHPAIIESQGTLEEVQDS 105
                                                                                                                   -----ERMTLENDDKTYTTEETLQVCE----QEDVPFVFDFHHFYANPDDHADL 225
                                                                                                                                    156 PESSQYPAVELEALENDNSD----SDTLDSSEEGGLEEEVA---RYEEKRYHPDRHRPL 207
                                                                                                                                                                                                                                                       : || | | | | | | :|| 327
                                                                                                                                                                                                                                                                                                                                                                                                                Q9UKX2; Q16229; Q14322; 20-AUG-2001 (Rel. 40, Created) 20-AUG-2001 (Rel. 40, Last sequence update) 20-AUG-2001 (Rel. 40, Last sequence update) MYOSIN HERVY CHAIN, SKELETAL MUSCLE, ADULT 2 (MYOSIN HERVY CHAIN, SKELETAL MUSCLE, ADULT 2 (MYOSIN HERVY CHAIN IIA)
93 FREIG-ELVKTHQLRTSFHPNQFTLFTSPKESVTKNAVTDMAYHYRMLEAMGI----- 144
                                                                                      106 MSETERS----ERMGARKRKDMSKK-----KGPPQEVKKGGEKEGSYHSPLNKSKKKKK 155
                                                                                                                                                                            226 ----NVALPRMIKTWERIGLQP----IR 257
                                                                                                                                                                                                        208 KTKMNVRPPPINPAGSRPSAPPPYELRLNTGTDSFLPLEERRKIQMAFPVFENAEGGRVH 267
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      "Type IIx myosin heavy chain transcripts are expressed in type IIb fibers of human skeletal muscle."; Am. J. Physiol. 267:C1723-C1728(1994).
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   Smerdu V., Karsch-Mizrachi I., Campione M., Leinwand L., Schiaffino S.;
                                                         ---ADRSVINIHIGGAYGNKDTATAQFHQNIKQLPQEIK----
                                                                                                                                                                                                                                                                                                                                                                                                     PRT; 1941 AA.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              TISSUE=Skeletal muscle;
MEDLINE=95270723; PubMed=7751403;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             MEDLINE=95109625; PubMed=7545970;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              SEQUENCE OF 1711-1941 FROM N.A.
TISSUE=Skeletal muscle;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                SEQUENCE OF 1823-1941 FROM N.A.
                                                                                                                                                                                                                                                                                                                                                                                                      STANDARD;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       Homo sapiens (Human).
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          SEQUENCE FROM N.A.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                NCBI_TaxID=9606;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       MYH2 OR MYHSA2.
                                                                                                                                                                                                                                                                                               S 317
                                                                                                                                                                                                                                                                                                                           328 A 328
                                                                                                                                                                                                                                                                                                                                                                                                   MYH2_HUMAN
                                                                                                                                                                                                                                                                                               317
                                                                                  q
                                                                                                                                               g
                                                                                                                                                                                                      g
                                                                                                                  ŏ
                                                                                                                                                                          ÓΥ
                                                                                                                                                                                                                                     δ
                                                                                                                                                                                                                                                              g
                                                                                                                                                                                                                                                                                                                          g
                                                                                                                                                                                                                                                                                                                                                                                                     δ
```

```
This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation - the Enropean Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See http://www.isb-sib.ch/announce/or send an email to license@lsb-sib.ch).
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     15;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               966 TLAKVEKEKHATENKVKNLTEEMAGLDETLAKLTKEKKALQEAHQQTLD------ 1014
-!- PTM: TWO CYSTEINE RESIDUES IN THE SI DOMAIN ARE SELECTIVELY ALKYLATED AND ARE REQUIRED FOR MYOSIN ATPASE ACTIVITY.
-!- MISCELLANEOUS: EACH MYOSIN HEAVY CHAIN CAN BE SPLIT INTO I LIGHT
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     111 PNQFTLFTSPKESVTKNAVTDMAYHYRMLEAMGIADRSVINIHIGGAYGNKDTATAQFHQ 170
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     52 MRTLHYIIGHGIPLYRFSSSIVPLATHPDVMWDFVTPFQKEFREI-GELVKTHQLRTSFH 110
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   828 MNVKHW-----PWMKLFFKIKPLLKSAETEKEMAT-MKEEFQKIKDELAKSEAKRKELE 880
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         881 EKMVTLL---KE---KN----DLQLQVQ-AEAEGLAD--------AEERCDQ 913
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           228 ALPRMIK------TWERIGLQPKVHLSSPKSEQAIRSHADYVDANFLLERFRQW 275
                                                          MEROMYOSIN (LMM) AND 1 HEAVY MEROMYOSIN (HMM), IT CAN LATER BE SPLIT FURTHER INTO 2 GLOBULAR SUBFRAGMENTS (S1) AND 1 ROD-SHAPED
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         769 VFFKAGLLGLLEEMRDDKLAQLITRTQARCRGFLARVEYQRMVERREAIFCI-QYNIRSF 827
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 171 NIK---QLPQEIKERMTLENDDKTYTTEETLQVCEQEDVPFVFDFHHFYANPDDHADLNV 227
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     Gaps
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   1 MIFREGEVSNAMSLWDASPAKTLT------FARYSKLSKTERKEALLTVTKANLRNT 51
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              Myosin; Muscle protein; Coiled coil; Thick filament; Actin-binding; Calmodulin-binding; ATP-binding; Methylation; Alkylation;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   86;
                                                                                                                -!- SIMILARITY: CONTAINS 1 MYOSTN-LIKE GLOBULAR HEAD DOMAIN. -!- SIMILARITY: CONTAINS 1 1Q DOMAIN.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             5.4%; Score 90; DB 1; Length 1941;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               46; Mismatches 130; Indels
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             ATP (POTENTIAL).
K -> R (IN REF. 3).
MW; 681E866F83AEA83F CRC64;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               COILED COIL (POTENTIAL).
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           MYOSIN HEAD-LIKE.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 Pred. No. 36;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        1015 ---- DLQAEEDKVNTLTKAKIKLEQQVDDLEG 1042
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   276 GTNIDFMIEAKQKD---KALLRLMDELSSIRG 304
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     767 816 10.
845 1941 COLLE
179 186 ATP (
1844 1844 1845 1941 AA; 223043 MW; 6
                                                                                                                                                                                                                                                                                                                                                                                                                                          InterPro; IPR()02928; Myosin_tail.
InterPro; IPR()01609; myosin_head.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              Pfam; PF00063; myosin_head; 1.
Pfam; PF01576; Myosin_tail; 1.
PRINTS; PR00193; MYOSINHEAVY.
ProDom; PD000355; myosin_head; 1.
                                                                                                                                                                                                                                                                                                                            EMBL; AF111784; AAD29950.1; -.
                                                                                                                                                                                                                                                                                                                                          EMBL; S73840; AAC13916.1; -.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           21.18;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           Local Similarity 21.1 hes 70; Conservative
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        SMART; SM00015; IQ; 1.
SMART; SM00242; MYSc; 1.
PROSITE; PS50096; IQ; 1.
                                                                                                                                                                                                                                                                                                                                                                                                                      InterPro; IPR000048; IQ.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         786
816
1941
186
1844
                                                                                            SUBFRAGMENT (S2)
                                                                                                                                                                                                                                                                                                                                                                              HSSP; P08799; 1MND.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                              Pfam; PF00612; IQ;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         Multigene family.
                                                                                                                                                                                                                                                                                                                                                                                                  MIM; 160740;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           Query Match
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               CONFLICT
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   SEQUENCE
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               ONIB ON
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         DOMAIN
DOMAIN
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         DOMAIN
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             Matches
  õ
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 qq
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               ó
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           q
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   qq
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               ^{\circ}
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             Q
```

σ

```
between the Swiss Institute of Bioinformatics and the EMBL outstation the European Bioinformatics Institute. There are no restrictions on its most by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See http://www.isb-sib.ch/announce/or send an email to license@isb-sib.ch).
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             Skelton J., Churcher C.M., Barrell B.G., Rajandream M.A., Wood V.; Skelton (AUG-1997) to the EMBL/GenBank/DDBJ databases.
-!- FUNCTION: STABILIZES THE F-ACTIN CABLES FORMING THE F-ACTIN RING THAT SURROUNDS THE NUCLEUS DURING INTERPHASE. MAY WORK IN
                                                                                                                                                                                                                                                                                                                                                                                                     MEDITINE-98062346; PubMed=9398685; Bezanilla M., Forsburg S.L., Pollard T.D.; Internstition of a second myosin-II in Schizosaccharomyces pombe: "Identification of a second myosin-II in Schizosaccharomyces pombe: Myp2p is conditionally required for cytokinesis."; Mol. Biol. Cell 8:2693-2705(1997).
                                                                                                                                                                                                                                                                              MEDLINE-90119390; PubMed-9459302; MEDLINE-90119390; Nakano K., Kitayama C., Yamamoto M., Mabuchi I.; Motegi F., Nakano K., Kitayama C., Yamamoto M., Mabuchi I.; Indentification of Myo3, a second type-II myosin heavy chain in the fission yeast Schizosaccharomyces pombe.";
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   ATP-binding; Actin-binding; Alkylation.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            -:- SUBGNIT: BINDS TO CDC4 AND RLC1.
-:- SIMILARITY: CONTAINS 1 MYOSIN-LIKE GLOBULAR HEAD DOMAIN.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       D -> G (IN REF. 2).
E -> G (IN REF. 2).
E -> K (IN REF. 2).
G -> D (IN REF. 2).
AM: 3AO548594028D258 CRC64;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       ATP (POTENTIAL).
ALKYLATION (BY SIMILARITY).
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       ACTIN-BINDING (POTENTIAL).
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    COILED COIL (POTENTIAL)
                                                                                                                                                      Schizosaccharomyces pombe (Fission yeast).
Eukaryota; Fungi; Ascomycota; Schizosaccharomycetes;
Schizosaccharomycetales; Schizosaccharomycetaceae;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   MYOSIN HEAD-LIKE
                               MYS3_SCHPO STANDARD; PRT; 2104 AA. 014157; 042730; 20-AUG-2001 (Rel. 40, Created) 20-AUG-2001 (Rel. 40, Last sequence update) 20-AUG-2001 (Rel. 40, Last annotation update) MYOSIN TYPE II HEAVY CHAIN 2.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            Pfan; PF00063; myosin_head; 1.
PRINTS; PR00193; MYOSINHEAVY.
ProDom; PD0000355; myosin_head; 1.
SMART; SM00242; MYSC; 1.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             IPR001609; myosin_head.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    EMBL; AB007633; BAA24579.1; -. EMBL; AF029788; AAC04615.1; -. EMBL; 298762; CAB11475.1; -. HSSP; P08799; IMND.
                                                                                                                                                                                                                                                                                                                                                    FEBS Lett. 420:161-166(1997).
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  242570
                                                                                                                                       MYO3 OR MYP2 OR SPAC4A8.05C
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             CONJUNCTION WITH MYO2
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 1420
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  AA;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      Myosin; Coiled coil;
                                                                                                                                                                                                                Schizosaccharomyces.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     SEQUENCE FROM N.A.
                                                                                                                                                                                                                                                                 SEQUENCE FROM N.A.
                                                                                                                                                                                                                                                                                                                                                                                           SEQUENCE FROM N.A.
                                                                                                                                                                                                                             NCBI_TaxID=4896;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         STRAIN=972;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 InterPro;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       DOMAIN
NP_BIND
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   CONFLICT
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  SEQUENCE
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       DOMAIN
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            MOD RES
RESULT 12
MYS3_SCHPO
```

12;

Gaps

58; Mismatches 130; Indels 80;

DB 1; Length 2104;

5.4%; Score 90; 19.8%; Pred. No.

Best Local Similarity 19.8%;

Query Match

66; Conservative

Matches

```
This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See http://www.isb-sib.ch/announce/
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             HELICAL, AND MANY OF THESE HELICES ARE AMPHIPATHIC. THEY MAY THEREFORE SERVE AS LIPID-BINDING DOMAINS WITH LECITHIN:CHOLESTEROL ACYLTRANSFERASE (LCAT) ACTIVATING ABILITIES.
SIMILARITY: BELONGS TO THE APOAl / APOA / APOE FAMILY.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 22-MER IS ACTUALLY A TANDEM ARRAY OF TWO, A AND B, RELATED 11-MERS) OCCURRING IN THIS SEQUENCE ARE PREDICTED TO BE HIGHLY ALPHA-
                                                                                                                                                                                                                                                                                                                                                                          : | : | : | : | : | : | : | | : | : | | : | : | : | | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | 
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            1081 DQLHKAFQNPEKTHNINDVKSGPLNSDENIYSTSSTTLSILKDVQELKSLHTK--EANOL 1138
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   Osada J., Pocovi M., Nicolosi R.J., Schaefer E.J., Ordovas J.M.;
"Nucleotide sequences of the Macaca fascicularis apolipoprotein C-III
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               and A-TV genes.",
Biochim. Biophys. Acta 1172:335-339(1993).
Biochim. Biophys. Acta 1172:335-339(1993).
-!- FUNCTION: MAY HAVE A ROLE IN CHYLOMICRONS AND VLDL SECRETION AND CATABOLISM. REQUIRED FOR EFFICIENT ACTIVATION OF LIPOPROTEIN LIPASE BY APOC-11; POTENT ACTIVATOR OF LCAT. APOA-IV IS A MAJOR COMPONENT OF HDL AND CHYLOMICRONS.
-!- SUBCELLULAR LOCATION: EXTRACELLULAR.
-!- TISSUE SPECIFICITY: SECRETED IN PLASMA.
-!- DOMAIN: NINE OF THE THIRTEEN 22-AMINO ACID TANDEM REPEATS (EACH
                                                                                863 TITKERVLOLTNDLQEEQALAHEKDILVERANSR--VEVVHERL-----SSLENQ 910
                                                                                                                                                                76 ATHPDVMMDEVTPFQKEFREIGELVKTHQLRTSFHPNQFTLFTSPKESVTKNAVTDMAYH 135
                                                                                                                                                                                                                                                    911 VTIADEKYEFL---YAEKQSIEEDLANKQTEISYLSDLSSTLEKKLSSIKKDEQT-ISSK 966
22 TLTFARYSKLSKTERKEALLT-----VTKANLRNTMRTLHYIIGHGIPLYRFSSSIVPL 75
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        210 FDFHHFYANPDDHADLNVALPRMIKTWERIGLQPKVHLSSPKSEQAIRS-HADYVDANFL
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Cercopithecidae;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            180 ------KERMTLENDDKTYTT-------EETLQVCEQEDVPFV
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         Macaca fascicularis (Crab eating macaque) (Cynomolgus monkey).
                                                                                                                                                                                                                                                                                                                                136 YRMLEAMGTADRSVINI----HIGGAYGNKDTATAOFHONIKQLPQEI----
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 01-FEB-1994 (Rel. 28, Last sequence update) .
15-DEC-1998 (Rel. 37, Last annotation update)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    1139 SERIKEISEMLEQSIATEEK---LRRKNSELCDI 1169
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   AA.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        269 LERFRQWGTNIDFMIEAKOKDKALLRLMDELSSI 302
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        or send an email to license@isb-sib.ch).
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 APOLIPOPROTEIN A-IV PRECURSOR (APO-AIV).
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       TISSUE=Leukocyte;
MEDLINE=93192330; PubMed=8448212;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   01-FEB-1994 (Rel. 28, Created)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               EMBL, X68361; CAA48421.1; -. P1R; S29565; S29565. PIR; S30195; S30195. HSSP; P02649; 1NFO.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   STANDARD;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                Cercopithecinae; Macaca.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 SEQUENCE FROM N.A.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 NCBI_TaxID=9541;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               15-DEC-1998
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       APA4 MACFA
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  P33621
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    APOA4
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             APA4_MACFA
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         Dp
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     qq
                                                                                                                                                                                     QΫ́
```

```
13;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  15-JUL-1999 (Rel. 38, Created)
15-JUL-1999 (Rel. 38, Last sequence update)
20-JUL-1999 (Rel. 30, Last annotation update)
20-AUG-2001 (Rel. 40, Last annotation update)
BACITRACIN SYNTHETASE (18A1) [INCLUDES: ATP-DEPENDENT ISOLEUCINE
ADENTLASE (ILRA) (ISOLEUCINE ACTIVASE); ATP-DEPENDENT CYSTEINE
ADENTLASE (CYSA) (CYSTEINE ACTIVASE); ATP-DEPENDENT LEUCINE ADENTLASE
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       --FSSSIVPLAT--HPDVMMD---FVTPFQKEFRE-------IGELVKTHQL 105
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     74 GDLOKKLVPFATELHERLAKDSEKLKEEIRKELEEVRARLLPHANEVSOKIGENVRELQO 133
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   106 RTSFHPNQFTLFTSPKESVTKNAVTDMAYHY-RMLEAMGIADRSVINIHIGGAYGNKDTA 164
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 134 RLEPYTDQLRTQVNTQTEQLRRQLTPYAQRMERVLRENADSLQTSLRPH-----ADQL 186
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             187 KAKIDONY----EELKERLTPYADEFKVKIDOTVEELRRSLAPYA------QDAQEK 233
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 165 TAQFHQNIKQLPQEIKERMTLENDDKTYTTETLQVCEQEDVPFVFDFHHFYANPDDHAD 224
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             225 LNVALP----RMIKTWERIGLQPKVHLSSPKSEQAIRSHADYVDANF-----LLERFRQ 274
                                                                                                                                                                                                                                                                                                                                                                                                                                                        15 VTGARAEVSADQVATVMWDYFSQLS-SNAKEAVEHLQKSELTQQLNALFQDKLGEVNTYA 73
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           234 LNHQLEGLAFQMKKNAEE--LKARISASAEELRQRLAPLAEDMRGNLRGNTEGLQKSLAE 291
                                                                                                                                                                                                                                                                                                                                                                                                      Gaps
                                                                                                                                                                                                                                                                                                                                                                                                                          8 VSNAMSLWDASPAKTLTFARYSKLSKTERKEALLTVTKANLRNTMRTLHYIIGHGIPLYR 67
                                                                                 13 X 22 AA APPROXIMATE TANDEM REPEATS.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             (LEUA) (LEUCINE ACTIVASE); ATP-DEPENDENT GLUTAMATE ADENYLASE (GLUA) (GLUTAMATE ACTIVASE); ATP-DEPENDENT ISOLEUCINE ADENYLASE (ILEA) (ISOLEUCINE ACTIVASE); GLUTAMATE RACEMASE (EC 5.1.1.3)].
                                                                                                                                                                                                                                                                                                                                                                                                    65;
                                                                                                                                                                                                                                                                                                                                                                   5.4%; Score 89.5; DB 1; Length 429; 20.4%; Pred. No. 5.2;
                              Chylomicron; Repeat; Signal
                                                                                                                                                                                                                                                                                                                                                                                  ; Pred. No. 5.2;
59; Mismatches 142; Indels
                                                                                                                                                                                                                                                                                                                      3D458F551D0DB60C CRC64;
                                                  BY SIMILARITY.
APOLIPOPROTEIN A-IV.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             Bacteria; Firmicutes; Bacillus/Clostridium group;
Bacillus/Staphylococcus group; Bacillus.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            PRT; 5255 AA.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            275 WGTNIDFMIE------AKQKDKALLRLMDEL 299
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       292 LGGHLDRHVEEFRLRVEPYGENFNKALVQQMEQL 325
                                                                                                                                                                                                                                                                                                       GLU/GLN-RICH
InterPro; IPR000074; Apolipoprotein.
                Pfam; PF01442; Apolipoprotein;
                                 Plasma; Lipid transport; HDL;
                                                                                                                                                                                                                                                                                                                    49876 MW;
                                                                                                                                                                                                                                                                                                                                                     Ouery Match
Best Local Similarity 20.4%
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         STANDARD;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              Bacillus licheniformis.
                                                                                                                                                                                                                                                                                                                  429 AA;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         SEQUENCE FROM N.A.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          STRAIN=ATCC 10716;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             NCBI_TaxID=1402;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       BACA_BACLI
068006;
                                                                                                                                                                                                                                                                                                                 SEQUENCE
                                                                                                                           REPEAT
REPEAT
                                                                                                                                                                                                                       REPEAT
REPEAT
                                                                                DOMAIN
                                                                                             REPEAT
                                                                                                              REPEAT
                                                                                                                                                            REPEAT
                                                                                                                                                                                         REPEAT
                                                                                                                                                                                                         REPEAT
                                                                                                                                                                                                                                                                                    REPEAT
                                                  SIGNAL
                                                                                                                                                                           REPEAT
                                                                                                                                                                                                                                                      REPEAT
                                                                                                                                                                                                                                                                     REPEAT
                                                                                                                                                                                                                                                                                                   DOMAIN
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         89
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         RESULT 14
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        BACA_BACLI
   qq
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      Q
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     ŏ
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                Q
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             δ
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           ò
```

```
ANTHRAY: NON-KIBOSOMAL BIOSYNTHESIS OF THE CYCLIC PEPTIDE
ANTIBIOTIC BACITRACIN.

SUGURIT: LARGE MULTENZYME COMPLEX OF BAI, BAZ AND BA3.

-!- SUGURIT: LARGE MULTENZYME COMPLEX OF BAI, BAZ AND BA3.

-!- DOMAIN: CONSIGNS OF FIVE MODULE INCORPORATES ONE ANINO ACID INTO THE FOURTH MODULE. EACH MODULE INCORPORATES ONE ANINO ACID INTO COMPANISH OF THE INITIATION MODULE), AND EPIMERIZATION CONDENSATION (NOT FOR THE INITIATION MODULE), AND EPIMERIZATION (OPTIONAL), AND NETHYATION (OPTIONAL).

-!- MISCELLANBOUS: BACITRACIN IS A MIXTURE OF AT LEAST TEN CYCLIC DODECAPEPTIDES, THAT DIFFER BY ONE OR TWO AMINO ACIDS, THE MOST CONTAINS AN N-TERRINAL LINEAR PENTAPEPTIDE MOIETY (ILE-CYS-LEU-D-GLU-ILE) WITH AN ISOLEUCINE-CYSTEINE THIAZOLINE CONDENSATION CONTAINS AND ACTERMINAL HEPTREPPIDE RING (LYS-D-ORF-ILE-D-PHE-HIS-D-ASP-ASN), IN WHICH THE FREE ALPHA-CARBOXY GROUP OF THE CONTAINS FANINO ACIDS IN THE D-CONFIGURATION (GLU-4, ORN-7, NINE) AND AND ACIDS IN THE D-CONFIGURATION (GLU-4, ORN-7),
                                                                                                                                                                                                     chem. Biol. 4:927-937(1997).
-1- FUNCTION: ACTIVATES FIVE AMINO ACIDS, INCORPORATES TWO D-AMINO ACIDS, RELEASES AND CYCLIZES THE MATURE BACITRACIN.
                                    Konz D., Klens A., Schoergendorfer K., Marahiel M.A.; "The bacitracin biosynthesis operon of Bacillus licheniformis ATCC
                                                                                                                            10716: molecular characterization of three multi-modular peptide
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  PHE-9, AND ASP-11). SIMILARITY: BELONGS TO THE ATP-DEPENDENT AMP-BINDING ENZYME
                                                                                                                                                                                                                                                                                                                                                    -!- CATALYTIC ACTIVITY: L-GLUTAMATE = D-GLUTAMATE.
-!- COFACTOR: CONTAINS 5 COVALENTLY BOUND PHOSPHOPANTETHEINES
MEDLINE=98089193; PubMed=9427658;
                                                                                                                                                                                                                                                                                                                                                                                                                                           (POTENTIAL)
                                                                                                                                                                              synthetases."
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              FAMILY
```

```
This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation - the Buropean Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See http://www.isb-sib.ch/announce/or send an email to license@isb-sib.ch).
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         c SIMILARITY).
f SIMILARITY).
f SIMILARITY).
f SIMILARITY).
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           Pfan: PF00501; AMP-binding; 5.
Pfan: PF00505; pp-binding; 5.
Pfan: PF00505; pp-binding; 5.
PR (NTS: PR0014; AMP-BINDING; 5.
PROSTTE; PS004012; PHOSPHOPANTETHEINE; 3.
PROSTTE; PS00455; AMP-BINDING; 5.
PROSTTE; PS00455; AMP-BINDING; 5.
Liqase; Isomerase; Antibiotic biosynthesis; Phosphopantetheine;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               PHOSPHOPANTETHEINE (BY SIMILARITY).
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     DOMAIN 1 (1SOLEUCINE-ACTIVATING).
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  DOMAIN 5 (ISOLEUCINE-ACTIVATING).
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               DOMAIN 2 (CYSTEINE-ACTIVATING).
DOMAIN 3 (LEUCINE-ACTIVATING).
DOMAIN 4 (GLUTAMINE-ACTIVATING).
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                PHOPANTETHEINE (BY SIMII)
PHOPANTETHEINE (BY SIMII)
906E8DD68450F85B CRC64;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         CYCLIZATION (POTENTIAL).
ACYL CARRIER (ACP) 1.
ACYL CARRIER (ACP) 2.
ACYL CARRIER (ACP) 3.
ACYL CARRIER (ACP) 4.
ACYL CARRIER (ACP) 5.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          PHOSPHOPANTETHEINE
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    PHOSPHOPANTETHEINE
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 PHOSPHOPANTETHEINE
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          PHOSPHOPANTETHEINE
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    InterPro; IPR003880; Phosphopant_attach.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        Multifunctional enzyme; Repeat.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                              InterPro; IPR000873; AMP-bind.
InterPro; IPR001242; DUF4.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                   EMBL; AF007865; AAC06346.1; -
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           2689
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           1037
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               1109
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             4668
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             3694
5201
5255
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          2124
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    544
1585
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   3664
5171
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      1615
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             BINDING
BINDING
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             REPEAT
DOMAIN
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             BINDING
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      BINDING
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    BINDING
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    REPEAT
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               REPEAT
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          REPEAT
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    REPEAT
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     DOMAIN
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  DOMAIN
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 DOMAIN
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            DOMAIN
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      DOMAIN
```

598254 MW;

AA;

SEQUENCE

```
Search completed: January 15, 2002, 14:07:20 Job time: 949 sec
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          InterPro; IPR002309; tRNA-synt_2.
InterPro; IPR002312; tRNA-synt_asp.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    Pfam; PF00152; tRNA-synt_2; 1. Pfam; PF01336; tRNA_anti; 1.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          550 AA; 64237 MW;
                                                                                                                                                                                                                                                                                                                                            EMBL; U39683; AAC71252.1; -. EEMBL; U01814; AAD12251.1; -. EMBL; X61523; CAA43736.1; -. TIGR; MG036; -.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        Conservative
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              470
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   377
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       COMPLICT 374
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      Local Similarity
nes 63; Conserv
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              469
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          SEQUENCE
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   Query Match
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 CONFLICT
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              Matches
SOFFE WAR WAR BURN SOFFE SOFFF SOFFE SOFFF SOFF 
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            Q
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           ŏ
                                                       14;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            2960 VGTPAAGRNHEDIQHLIGMFVNTLAIRNHPEGKKTFRDYLQEVKENTLQAYENQDYPFEE 3019
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   PETERSON S.N., Schramm N., Hu P.-C., Bott K.F., Hutchison C.A. III;
Peterson S.N., Schramm N., Hu P.-C., Bott K.F., Hutchison C.A. III;
"A random sequencing approach for placing markers on the physical map of Mycoplasma genitalium.";
Nucleic Anids Res. 19:6027-6031(1991).
-! CATALYTIC ACTIVITY: ATP + L-ASPARTATE + TRNA(ASP) = AMP + PYROPHOSPHATE + L-ASPARTYL-TRNA(ASP).
                                                                                                                                                2789 ESAEEKIKSFIRPFHLEKAPLLRA--EVVKLNEREHLLMFDMHHIISDGVSTDIFIQELG 2846
                                                                                                                                                                                                                                                                                                                                                                                                             2900 DHKRPLTKSSEGDTVTAAIESETFRKLQHMAKENGVTMYMLLLAGYTALLSKYTGQEDII 2959
                                                                                                                                                                                                                                                                                           2847 ALYEGKS-----LKPFHIQYKDYAEWENSHARSEELKRQEEYWLKT--YKGDIPVLDLPI 2899
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               Fraser C.M., Gocayne J.D., White O., Adams M.D., Clayton R.A., Fleischmann R.D., Bult C.J., Kerlavage A.R., Sutton G., Kelley J.M., Frieischmann R.D., Bult C.J., Kerlavage A.R., Sutton G., Kelley J.M., Fritchman J.L., Weidman J.F., Small K.V., Sandusky M., Fuhrmann J.L., Nguyen D.T., Utterback T.R., Saudek D.M., Phillips C.A., Merrick J.M., Pomb J.-F., Dougherty B.A., Bott K.F., Hu P.-C., Lucier T.S., Peterson S.N., Smith H.O., Hutchison C.A. III, Venter J.C.; "The minimal gene complement of Mycoplasma genitalium.";
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 3020 LVEKVNIKRDMARNPLFDTMLVYHNTD------VKPFEAEGLRSRLVEIKRGIS 3067
                                                                                                                                                                                                                                74 PLATHPDVMWDFVTPFQKEFREIGELVKTHQ------LRTSFHPNQFTLFTSP- 120
                                                                                                                                                                                                                                                                                                                                                                                                                                                                    ---TATAQFHQNIKQLPQEIKERMTLEN---DDKTY-----TTEETLQVCEQEDVPF-- 208
                                                          Gaps
                                                                                                              16 DASPAKTLTFARYSKLSKTERKEALLTVTKANLRN--TMRTLHYIIGHGIPLYRFSSSIV 73
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         ------VFDFHHFYANPDDHADLNVALPRMIKTWERIGLQPK-----
                                                                                                                                                                                                                                                                                                                                                     --KESVTKNAVTDM-----AYHYRMLEAMGIADRSVINIHIGGAY------GNKD--
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   SYD_MYCGE STANDARD; PRT; 550 AA.
P47282; 049462; 049479;
01-FEB-1996 (Rel. 33, Created)
01-FEB-1996 (Rel. 33, Last sequence update)
20-AUG-2001 (Rel. 40, Last annotation update)
ASPARTYL-TRNA SYNTHETASE (EC 6.1.1.12) (ASPARTATE--TRNA LIGASE)
                                                    95;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   STRAIN=ATCC 31530 / G-37;
MEDLINE=94075230; PubMed=8253680;
Peterson S.N., Hu P.-C., Bott K.F., Hutchison C.A. III;
"A survey of the Mycoplasma genitalium genome by using random
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       Mycoplasma genitalium.
Bacteria; Firmicutes; Bacillus/Clostridium group; Mollicutes:
5.4%; Score 89.5; DB 1; Length 5255; 18.7%; Pred. No. 1.5e+02;
                                                          Indels
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      3068 KFDITVTASEAADGLRLEVEYSTTLFNKERMERLSEHLISLLE 3110
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              245 ---VHLSSPKSEQAIRSHADYVDANFLLERFRQWGTNIDFMIE 284
                                                       54; Mismatches 130;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  SEQUENCE FROM N.A.
STRAIN-ATCC 33530 / G-37;
MEDLINE-96026346; PubMed-7569993;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             Bacteriol, 175:7918-7930(1993)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           SEQUENCE OF 374-550 FROM N.A. STRAIN=ATCC 33530 / G-37;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  SEQUENCE OF 388-470 FROM N.A.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  Mycoplasmataceae; Mycoplasma.
                                                          64; Conservative
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   STRAIN=ATCC 33530 / G-37
Query Match
Best Local Similarity
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             NCBI_TaxID=2097;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           ASPS OR MG036.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     sequencing
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      (ASPRS)
                                                                                                                                                                                                                                                                                                                                                        121
                                                             Matches
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               209
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        q
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      qq
                                                                                                                                                                                                                                                                                                                                                                                                                g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 ð
                                                                                                                                                                                                                                                                                                                                                     ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                          ò
                                                                                                                     ò
                                                                                                                                                                                                                                          ò
```

```
This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See http://www.isb-sib.ch/announce/or send an email to license@isb-sib.ch).
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  16;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    333 HDVSLEVHNKNNSTLKTSIKCDEKNTL------LLVANKSKKKAWTALG 375
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               -!- SUBUNIT: HOMODIMER (BY SIMILARITY).
-!- SUBCELLULAR LOCATION: CYTOPLASMIC.
-!- SIMILARITY: BELONGS TO CLASS-1I AMINOACYL-TRNA SYNTHETASE FAMILY.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            64 PLYRFSSSIVPLAT-HPDVMWDFVTPFQKEFREIGELVKTHQLRTSFHPNQFTLFTSPKE 122
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   123 SVTKNAVTDMAYHYRMLEAMGIADRSVINIHIGG-AYGNKDTATAQFHQNIKQLPQEIKE 181
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           Aminoacyl-tRNA synthetase; Protein biosynthesis; Ligase; ATP-binding;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      Gaps
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                300 ------KLIEAKGI-----IGGIELTNIELST----DKIKALRKIAKD
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  241 L---OPKVHLSSPKSEQAIRSHADYVDANFLL----ERFRQWGTNIDFMIEAKQ-----
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          260 PVFKFSELFERFGSDKPDLRYGF-----EIKDFTSLFQDHQ-----NQFT-----
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      182 R-MILENDOKTYTTEETLQVCEQEDVPFVFDFHHFYANPDDHADLNVALPRMIKTWERIG
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      :06
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     5.4%; Score 89; DB 1; Length 550; 23.3%; Pred. No. 7.9;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      78; Indels
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      MN -> II (IN REF. 3).
4B7972691E38D4BA CRC64;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 LGAI -> CWSN (IN REF. 2).
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         PRINTS; PRO1042; TRNASYNTHASP.
PROSITE; PS00179; AA_TRNA_LIGASE_II_1; FALSE_NEG.
PROSITE; PS00339; AA_TRNA_LIGASE_II_2; 1.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  39; Mismatches
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             288 ---KDKALLRLMDELSSIRGVKRIGGGALQ 314
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       432 ESNKNEALSEQFDLV -- LNGFE-IGSGSIR 458
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 InterPro; IPR002106; AA_tRNA_ligase_II.
```

. .

```
GenCore version 4.5
Copyright (c) 1993 - 2000 Compugen Ltd.
```

OM protein - protein search, using sw model

January 15, 2002, 13:51:31; Search time 45.07 Seconds (without alignments) 673.585 Million cell updates/sec Run on:

Title: Perfect score:

US-09-724-296-2 4349 1 WTKLPILGYWKIKGLVQPTR......KRLTARKRRSRKEEVEEDEK 828 Sequence:

Scoring table:

100059 seqs, 36664827 residues BLOSUM62 Gapop 10.0 , Gapext 0.5 Searched:

100059 Total number of hits satisfying chosen parameters:

Minimum DB seq length: 0 Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

SwissProt_39:* Database :

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

STIMMARTES

Ouery Match Length 26.9 22.6 22.6 22.6 22.8 22.6 22.8 16.3 22.0 16.3 22.0 16.3 22.0 10.9 217 217 217 217 217 217 217 217 217 217	:
--	---

04906 rattus norv	P46424 cricetulus	28514 macaca mula	46425 mus musculu	28801 bos taurus	08863 oryctolagus	10299 caenorhabdi	04903 rattus norv	10648 mus musculu	13745 mus musculu	00502 rattus norv	46418 rattus norv
0 d	P4	02	P4	P2	00	P1	0d	P1	P1	P0	b4
GTP_RAT	GTP_CRILO	GTP_MACMU	GTP1_MOUSE	GTP_BOVIN	GTA1_RABIT	GTP_CAEEL	GTA2_RAT	GTA2_MOUSE	GTA1_MOUSE	GTA1_RAT	GTC2_RAT
П	-	,	~	_	1	_	_	,	Г	_	~
209	509	209	509	209	223	208	221	221	222	221	220
		_	4	~		0	9	9	9	2	4
5.4	5.4	5.4	7	ν.	5.	Ŋ.	4	4	7	4	4
	236 5.4				-		•		·	•	

ALIGNMENTS

```
This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss institute of Bioinformatics and the EMBL outstation the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See http://www.isb-sib.ch/announce/or send an email to license@isb-sib.ch).
                                                                                                                                                                   0;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    SEQUENCE FROM N.A.
MEDLINE=92131046; PubMed=1775156;
Wright M.D., Harrison R.A., Melder A.M., Newport G.R., Mitchell G.F.;
"Another 26-kilodalton glutathione S-transferase of Schistosoma mansoni.";
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        63 VKLTQSMAIIRYIADKHNMLGGCPKERAEISMLEGAVLDIRYGVSRIAYSKDFETLKVDF 122
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          123 LSKLPEMLKMFEDRLCHKTYLNGDHVTHPDFMLYDALDVVLYMDPMCLDAFPKLVCFKKR 182
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   65 VKLTQSMAIIRYIADKHNMLGGCPKERAEISMLEGAVLDIRYGVSRIAYSKDFETLKVDF 124
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     125 LSKLPEMLKMFEDRLCHKTYLNGDHVTHPDFMLYDALDVVLYMDPMCLDAFPKLVCFKKR 184
                                                                                                                                                                                                                                                                                                                                                                                                                                                                             Gaps
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          3 PILGYWKIKGLVQPTRLLLEYLEEKYEEHLYERDEGDKWRNKKFELGLEFPNLPYYIDGD 62
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    5 PILGYWKIKGLVOPTRLLLEYLEEKYEEHLYERDEGDKWRNKKFELGLEFPNLPYYIDGD 64
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   MOI. Biochem. Parasitol. 49:177-179(1991).

-!- FUNCTION: CONJUGATION OF REDUCED GLUTATHIONE TO A WIDE NUMBER OF EXOGENOUS AND ENDOGENOUS HYDROPHOBIC ELECTROPHILES.

-!- FUNCTION: GST ISORIZYMES APPEAR TO PLAY A CENTRAL ROLE IN THE PARASITE DEFOXIFICATION SYSTEM. OTHER FUNCTIONS ARE ALSO SUSPECTED INCLUDING A ROLE IN INCREASING THE SOLUBILITY OF HAEMATIN IN THE PARASITE GTT.

-!- CATALITIC ACTIVITY: RX + GLUTATHIONE = HX + R-S-GLUTATHIONE.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          Eukaryota; Metazoa; Platyhelminthes; Turbellarian Platyhelminths; Rhabditophora; Eulecithophora; Revertospermata; Mediofusata; Neodermata; Trematoda; Digenea; Strigeidida; Schistosomatidae; Schistosoma.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  01-JUN-1994 (Rel. 29, Created)
01-JUN-1994 (Rel. 29, Last sequence update)
01-NOV-1995 (Rel. 32, Last annotation update)
GLUTATHIONE S-TRANSFERASE 26 KDA (EC 2.5.1.18) (GST 26) (SM26/2
                                                                                                                                                                                                                                                                                                                                                                                                                                                                             .,
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         -!- SUBUNIT: HOMODIMER.
-!- SIMILARITY: BELONGS TO THE GST SUPERFAMILY. ALPHA FAMILY.
                                                                                                                                                                                                                                                                                                                                                                                                                                 26.9%; Score 1170; DB 1; Length 218; 100.0%; Pred. No. 2.8e-56;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                           Indels
                                                                                                                                                                                                                                                                                                                                                   Transferase; Antigen; Multigene family; 3D-structure. SEQUENCE 218 AA; 25498 MW; 5E2AC418BD0EF13F CRC64;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    185 IEAIPQIDKYLKSSKYIAWPLQGWQATFGGGDHPPK 220
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       218 AA.
                                                                                                                                                                                                                                                                                                                                                                                                                                            100.0%; Pred. ....
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 PRT;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        Schistosoma mansoni (Blood fluke).
                                                                                                                                                                                    EMBL; M14654; AAB59203.1; -.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     ANTIGEN) (GST CLASS-ALPHA).
                                                                                                                                                                                                                                                                                                        InterPro; IPR000521; GST.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                           Matches 216; Conservative
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               STANDARD;
                                                                                                                                                                                                                         1GNE; 30-NOV-94.
1GTA; 07-FEB-95.
1GTB; 01-DEC-95.
1B8X; 12-APR-99.
                                                                                                                                                                                                                                                                                                                             Pfam; PF00043; GST; 1.
                                                                                                                                                                                                        A26484; A26484.
                                                                                                                                                                                                                                                                                                                                                                                                                                    Query Match
Best Local Similarity
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               NCBI_TaxID=6183;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           GT27_SCHMA
P35661:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        RESULT 2
GT27_SCHMA
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          q
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    ŏ
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                δŏ
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     δ
```

```
This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation the Buropean Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way
                                                                                          modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See http://www.isb-sib.ch/announce/or send an email to license@isb-sib.ch).
                                                                                                                                                                                                                                                                                                                                                                                               0
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       65 VKLTQSMALIRYIADKHNMLGGCPKERAEISMLEGAVLDIRYGVSRIAYSKDFETLKVDF 124
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              125 LSKLPEMLKMFEDRLCHKTYLNGDHVTHPDFMLYDALDVVLYMDPMCLDAFPKLVCFKKR 184
                                                                                                                                                                                                                                                                                                                                                                                               0; Gaps
                                                                                                                                                                                                                                                                                                                                                                                                                                     5 PILGYWKIKGLVQPTRLLLEYLEEKYEEHLYERDEGDKWRNKKFELGLEFPNLPYYIDGD 64
                                                                                                                                                                                                                                                                                                                                                                                                                                                                          3 PKLGYWKIKGLVQPTRLLLEYLGERYEERLYDRNDGDVWRNEKFKLGLDFPNLPYYIDGD 62
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          glutathione S-transferases of Schistosoma japonicum and Schistosoma
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    Henkle K.J., Davern K.M., Wright M.D., Ramos A.J., Mitchell G.F., "Comparison of the cloned genes of the 26- and 28-kilodalton
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  MOI. Biochem. Parasitol. 40:23-34(1990).
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     ANTIGEN) (GST CLASS-ALPHA).
Schistosoma mansoni (Blood fluke).
Eukaryota; Metazoa; Platyhelminthes; Turbellarian Platyhelminths;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  Trottein F., Kieny M.P., Verwäerde C., Torpier G., Pierce R.J., Balloul J.-M., Schmitt D., Lecocq J.-P., Capron A.; "Molecular clonding and Lissue distribution of a 26-kilodalton Schistosoma mansoni glutathione S-transferase."; Mol. Biochem. Parasitol. 41:35-44(1990).
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            01-APR-1990 (Rel. 14, Last sequence update)
01-NOV-1995 (Rel. 32, Last annotation update)
GLUTATHIONE S-TRANSFERASE 26 KDA (EC 2.5.1.18) (GST 26) (SM26/1
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               Rhabditophora; Bulecithophora; Revertospermata; Mediofusata;
Neodermata; Trematoda; Digenea; Strigeidida; Schistosomatoidea;
                                                                                                                                                                                                                                                                                                                                                   22.6%; Score 984; DB 1; Length 218; 82.9%; Pred. No. 2.5e-46; Live 18; Mismatches 19; Indels
                                                                                                                                                                                                                                                   pfam: PF00043; GST; 1.
Transferase; Antigen; Multigene family.
SEOUENCE 218 AA; 25411 MW; D8D3EE9028B36185 CRC64;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             185 IEALPQIDKYLKSSKYIAWPLQGWQATFGGGDHPPK 220
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    218 AA.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      PRT;
                                                                                                                                                                       EMBL; M73624; -; NOT_ANNOTATED_CDS.
PIR; A45556; A45556.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        MEDLINE=90348716; PubMed=2385266;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   MEDLINE=90271935; PubMed=1693415;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         01-APR-1990 (Rel. 14, Created)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     Schistosomatidae; Schistosoma.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              SEQUENCE OF 8-218 FROM N.A.
                                                                                                                                                                                                                                                                                                                                                                                               Matches 179; Conservative
                                                                                                                                                                                                                                      InterPro; IPR000521; GST.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      STANDARD;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     STRAIN-PUERTO RICAN;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   STRAIN-PUERTO RICAN;
                                                                                                                                                                                                                                                                                                                                                                              Best Local Similarity
                                                                                                                                                                                                                 HSSP; P08515; 1GTB.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  SEQUENCE FROM N.A.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          NCBI_Tax1D=6183;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      GT26_SCHMA
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     mansoni."
                                                                                                                                                                                                                                                                                                                                                           Query Match
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               GT26_SCHMA
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      δ
                                                                                                                                                                                                                                                                                                                                                                                                                                                                             qq
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          οp
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      q
δ
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  δλ
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             ŏλ
```

```
Exp. Parasitol, 74:232-237(1992).
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   Fasciola hepatica (Liver fluke).
                                                                                                                                                                                                                                                                                                                                                                                                                                                                             22 22 110 111 DP 189 A 220 AA; 25598 MW;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               32;
                                                                                                                                                                                                                                                                                                                                                                                                          EMBL; A00993; CAA00118.1; -.
                                                                                                         SEQUENCE OF 22-220 FROM N.A.
                                                                                                                                                                                                                                                                                                                                                                                                                     InterPro, IPR000521; GST.
Pfam; PF00043; GST; 1.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               129; Conservative
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    STANDARD;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                   0
                                                                                                                                                    -! - FUNCTION:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  GT27_FASHE
P31670;
                                                                                                                        Crameri S.;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           CONFLICT
                                                                                                                                                                                                                                                                                                                                                                                                                                                                   INIT MET
                                                                                                                                                                                                                                                                                                                                                                                                                                                                            CONFLICT
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             CONFLICT
                                   ERRATUM.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      RESULT 5
GT27_FASHE
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              Matches
        qq
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            δy
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   Qy
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             q
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      q
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    ó
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  ô
                                                                                                                                                                                     This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation - the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See http://www.isb-sib.ch/announce/or send an email to license@isb-sib.ch).
                                                                                                                                                                                                                                                                                                                                                                                                                                                                               0
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            123 LNKLPGRLKMFEDRLSNKTYLNGNCVTHPDFMLYDALDVVLYMDSQCLNEFPKLVSFKKC 182
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             MEDINE=92155306; PubMed=1740183; Panaccio M., Wilson L.R., Crameri S.L., Wijffels G.L., Spithill T.W.; "Molecular characterization of cDNA sequences encoding glutathione Stransferases of Fasciola hepatica.";
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             65 VKLTQSMAIIRYIADKHNMLGGCPKERAEISMLEGAVLDIRYGVSRIAYSKDFETLKVDF 124
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      LSKLPEMLKMFEDRLCHKTYLNGDHVTHPDFMLYDALDVVLYMDPMCLDAFPKLVCFKKR 184
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       3 PKFGYWKVKGLVQPTRLLLEHLEETYEERAYDRNEIDAWSNDKFKLGLEFPNLPYYIDGD 62
                                                                                                                                                                                                                                                                                                                                                                                                                                                                               0; Gaps
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      5 PILGYWKIKGLVQPTRLLLEYLEEKYEEHLYERDEGDKWRNKKFELGLEFPNLPYYIDGD 64
OF EXOGENOUS AND ENDOCENOUS HYDROPHOBIC ELECTROPHILES.

-!- FUNCTION: GST ISOENXYMES APPEAR TO PLAY A CENTRAL ROLE IN THE PARASITE DETOXIFICATION SYSTEM. OTHER FUNCTIONS ARE ALSO SUSPECTED INCLUDING A ROLE IN INCREASING THE SOLUBILITY OF HAEMATIN IN THE PARASITE GUT.
-!- CATALYTIC ACTIVITY: RX + GLUTATHIONE = HX + R-S-GLUTATHIONE.
-!- SUBUNIT: HOMODIMER.
-!- SUBUNIT: HOMODIMER.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       15-DEC-1998 (Rel. 37, Last sequence update)
15-DEC-1998 (Rel. 37, Last annotation update)
GLUTATHIONE S-TRANSFERASE 26 KDA 1 (EC 2.5.1.18) (GST1) (FH1) (GST
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            Eukaryota; Metazoa; Platyhelminthes; Turbellarian Platyhelminths; Rhabditophora; Eulecithophora; Revertospermata; Mediofusata; Neodermata; Trematoda; Digenea; Echinostomida; Echinostomata; Pascioloidea; Fasciolaidae; Fasciola
                                                                                                                  CELLS. GST 26 MAY BE ACTIVELY EXCRETED BY ADULT WORMS.
-!- MISCELLANEOUS: THERE ARE AT LEAST THREE ISOENZYMES OF GST IN
                                                                                                                                                           -!- SIMILARITY: BELONGS TO THE GST SUPERFAMILY. ALPHA FAMILY.
                                                                                                                                                                                                                                                                                                                                                                                                                                               Length 218;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                             Indels
                                                                                                                                                                                                                                                                                                                                                                       Pfan, PP00043; GST; 1.
Transferase; Antigen; Multigene family.
SEQUENCE 218 AA; 25401 MW; 061A6548A842D6E8 CRC64;
                                                                                                                                                                                                                                                                                                                                                                                                                                              21.2%; Score 921; DB 1;
80.1%; Pred. No. 5.8e-43;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              185 IEAIPQIDKYLKSSKYIAWPLQGWQATFGGGDHPPK 220
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             183 IEDLPQIKNYLNSSRYIKWPLQGWDATFGGGDTPPK 218
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                220 AA.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                            14; Mismatches
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  PRT;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 Fasciola hepatica (Liver fluke).
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         15-DEC-1998 (Rel. 37, Created)
                                                                                                                                                                                                                                                                                                    EMBL; M31106; AAA29888.1; -.
                                                                                                                                                                                                                                                                                                                    EMBL; M26913; AAA29889.1; -.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                         Matches 173; Conservative
                                                                                                                                                                                                                                                                                                                                                        InterPro; IPR000521; GST.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                STANDARD;
                                                                                                                                                                                                                                                                                                                                PIR; A45523; A45523.
HSSP; P08515; 1GTB.
                                                                                                                                                                                                                                                                                                                                                                                                                                                              Similarity
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 SEQUENCE FROM N.A.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   NCBI_TaxID=6192;
                                                                                                                                               S. MANSONI
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  CLASS-ALPHA).
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             GT29_FASHE
P56598;
                                                                                                                                                                                                                                                                                                                                                                                                                                               Query Match
                                                                                                                                                                                                                                                                                                                                                                                                                                                              Best Local
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 GT29 FASHE
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      125
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    RESULT
    g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        qq
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                qq
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    δ
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             ŏ
```

```
This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation-the Buropean Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     entities requires a license agreement (See http://www.isb-sib.ch/announce/or send an email to license@isb-sib.ch).
                             Panaccio M., Wilson L.R., Crameri S.L., Wijffels G.L., Spithill T.W.;
Exp. Parasitol. 77:385-385(1993).
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        67 LTQSMAIIRYTADKHNMLGGCPKERAEISMLEGAVLDIRYGVSRIAYSKDFETLKVDFLS 126
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              127 KLPEMLKMFEDRLCHKTYLNGDHVTHPDFMLYDALDVVLYMDPMCLDAFPKLVCFKKRIE 186
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  Gaps
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 4 LGYWKIRGLQQPVRLLLEY-GEKYEEQIYERDDGEKWFSKKFELGLDLPNLPYYIDDKCK 62
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            7 LGYWKIKGLVOPTRLLLEYLEEKYEEHLYERDEGDKWRNKKFELGLEFPNLPYYIDGDVK 66
                                                                                                                                                                                                                                                                          OF EXOGENOUS AND ENDOGENOUS HYDROPHOBIC ELECTROPHILES. FUNCTION: GST ISOENZYMES APPEAR TO PLAY A CENTRAL ROLE IN THE PARASITE DETOXIEICATION SYSTEM. OTHER FUNCTIONS ARE ALSO
                                                                                                                                                                                                                                                CONJUGATION OF REDUCED GLUTATHIONE TO A WIDE NUMBER
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       Eukaryota; Metazoa; Platyhelminthes; Turbellarian Platyhelminths; Rhabditophora; Eulecithophora; Revertospermata; Mediofusata; Neodermata; Trematoda; Digenea; Echinostomida; Echinostomata;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             GLUTATHIONE S-TRANSFERASE 26 KDA 47 (EC 2.5.1.18) (GST47) (FH47) (GST CLASS-ALPHA).
                                                                                                                                                                                                                                                                                                                                                                                       HAEMATIN IN THE PARASITE GUT.
-!- CATALYTIC ACTIVITY: RX + GLUTATHIONE = HX + R-S-GLUTATHIONE.
-!- SUBUNIT: HOMODIMER.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  Query Match 16.3%; Score 710.5; DB 1; Length 220; Best Local Similarity 59.7%; Pred. No. 1.1e-31;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           -!- SIMILARITY: BELONGS TO THE GST SUPERFAMILY. ALPHA FAMILY
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           54; Indels
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           A -> P (IN REF. 3).
27B9F150B75D101F CRC64;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     BY SIMILARITY.
Y -> V (IN REF. 3).
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              (IN REF.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             187 AIPQIDKYLKSSKYIAWPLQGWQATFGGGDHPPKSD 222
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    01-JUL-1993 (Rel. 26, Created)
15-DEC-1998 (Rel. 37, Last sequence update)
15-DEC-1998 (Rel. 37, Last annotation update)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         217 AA
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           Mismatches
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                Transferase; Antigen; Multigene family
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          DP -> VS
                                                                                                                                                                                                        Patent number W09008819, 09-AUG-1990
MEDLINE=94039664; Pubmed=8224094;
```

ij

```
This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the PRBL outstation the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (see http://www.isb-sib.ch/announce/or send an email to license@isb-sib.ch).
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 ;
                                                                                                                                                                                                                    MEDLINE-94039664; PubMed-8224094; Panaccio M., Wilson L.R., Crameri S.L., Wilfels G.L., Spithill T.W.; Exp. Parasitol. 77:385-385(1993).
                                                                                             Panaccio M., Wilson L.R., Crameri S.L., Wijffels G.L., Spithill T.W.; "Molecular characterization of cDNA sequences encoding glutathione Stransferases of Fasciola hepatica.";
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        127 KLPEMLKMFEDRLCHKTYLNGDHVTHPDFMLYDALDVVLYMDPMCLDAFPKLVCFKKRIE 186
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            LTQSMAIIRYIADKHNMLGGCPKERAEISMLEGAVLDIRYGVSRIAYSKDFETLKVDFLS 126
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               64 LTQSVAIMRYIADKHGMLGTTPEERARISMIEGAAMDLRIGFGRVCYNPKFEEVKEEYVK 123
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 0; Gaps
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          7 LGYWKIKGLVQPTRLLLEYLEEKYEEHLYERDEGDKWRNKKFELGLEFPNLPYYIDGDVK 66
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               4 LGYWKLRGLAOPVRLFLEYLGEEYEEHLYGRDDREKWMSEKFNMGLDLPNLPYYIDDKCK 63
                                                                                                                                                                                                                                                                                                                                                                                                                                                                    "Crystallization, structural determination and analysis of a novel
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      OF EXOGENOUS AND ENDOGENOUS HYDROPHOBIC ELECTROPHILES.
-!- FUNCTION: GST ISOENZYMES APPEAR TO PLAY A CENTRAL ROLE IN THE PARASITE DETOXIFICATION SYSTEM. OTHER FUNCTIONS ARE ALSO
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             J. Mol. Biol. 273:857-872(1997).
-1- FUNCTION: CONJUGATION OF REDUCED GLUTATHIONE TO A WIDE NUMBER
                                                                                                                                                                                                                                                                                                                                                                                                          MEDLINE=98035725; PubMed=9367777;
Rossjohn J., Feil S.C., Wilce M.C.J., Sexton J.L., Spithill T.W.,
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              SUSPECTED INCLUDING A ROLE IN INCREASING THE SOLUBILITY OF HAEMATIN IN THE PARASITE GUT.

CATALYTIC ACTIVITY: RX + GLUTATHIONE = HX + R-S-GLUTATHIONE.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              -!- CATALLILA
-!- SUBUNIT: HOMODIMER.
-!- SUBCELLUAR LOCATION: CYTOPLASMIC.
-!- SIMILARITY: BELONGS TO THE GST SUPERFAMILY. ALPHA FAMILY.
-!- SIMILARITY: BELONGS TO THE GST SUPERFAMILY.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      Query Match 16.1%; Score 699; DB 1; Length 217; Best Local Similarity 58.2%; Pred. No. 4.3e-31; Matches 124; Conservative 36; Mismatches 53; Indels
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          parasite vaccine candidate: Fasciola hepatica glutathione
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  T -> I (IN REF. 3).
RIGF -> FEEL (IN REF. 3).
OFB8BBFE63029E03 CRC64;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             Transferase; Antigen; Multigene family; 3D-structure.
                                                                                                                                                                                                                                                                                                                                             Patent number W09008819, 09-AUG-1990.
                                                                                                                                                                                                                                                                                                                                                                                      X-RAY CRYSTALLOGRAPHY (3 ANGSTROMS).
Fascioloidea; Fasciolidae; Fasciola.
                                                          SEQUENCE FROM N.A.
MEDLINE=92155306; PubMed=1740183;
                                                                                                                                                            Exp. Parasitol, 74:232-237(1992).
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  65 65 T
102 105 R
217 AA; 25281 MW;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         EMBL; M77681; AAA29140.1; -. EMBL; A00996; CAA00121.1; -.
                                                                                                                                                                                                                                                                                                      SEQUENCE OF 7-105 FROM N.A.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                PDB; 1FHE; 29-JUL-98.
InterPro; IPR000521; GST.
Pfam; PF00043; GST; 1.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   0
               NCBI_TaxID=6192;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              S-transferase.";
                                                                                                                                                                                                                                                                                                                                                                                                                                                 Parker M.W.;
                                                                                                                                                                                                                                                                                                                              Crameri S.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        CONFLICT
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        CONFLICT
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     INIT_MET
                                                                                                                                                                                                       ERRATUM
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 q
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               g
οy
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          δ
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        δλ
```

```
This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the BMBL outstation the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See http://www.isb-sib.ch/announce/or send an email to license@lsb-sib.ch).
                                                                                                                                                                                                                                                                                                                                                                                                                                                       Panaccio M., Wilson L.R., Crameri S.L., Wijffels G.L., Spithill T.W., "Molecular characterization of cDNA sequences encoding glutathione S-transferases of Fasciola hepatica."; 
Exp. Parasitol. 74:232-237(1992).
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  Panaccio M., Wilson L.R., Crameri S.L., Wijffels G.L., Spithill T.W.;
Exp. Parasitol. 77:385-385(1993).
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              0; Gaps
                                                                                                                                                                                             [5-DEC-1998 (Rel. 37, Last sequence update)
15-DEC-1998 (Rel. 37, Last annotation update)
GLUTATHIONE S-TRANSFERASE 26 KDA 7 (EC 2.5.1.18) (GST) (FH7) (GST
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               Patent number W09008819, 09-AUG-1990.

-!- FUNCTION: CONJUGATION OF REDUCED GLUTATHIONE TO A WIDE NUMBER OF EXOGENOUS AND ENDOGENOUS HYDROPHOBIC ELECTROPHILES.

-!- FUNCTION: GST ISOBENZYMES APPEAR TO PLAY A CENTRAL ROLE IN THE PARASITE DETOXIFICATION SYSTEM. OTHER FUNCTIONS ARE ALSO
                                                                                                                                                                                                                                                                                                 Eukaryota; Metazoa; Platyhelminthes; Turbellarian Platyhelminths;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            SUSPECTED INCLUDING A ROLE IN INCREASING THE SOLUBILITY OF HAEMATIN IN THE PARASITE GUT.

CATALYTIC ACTIVITY: RX + GLUTATHIONE = HX + R-S-GLUTATHIONE.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     NDREKWLG -> MIGRNGWA (IN REF. 3).
IKEYMKSER -> SRYMSRA (IN REF. 3).
                                                                                                                                                                                                                                                                                                                    Rhabditophora; Eulecithophora; Revertospermata; Mediofusata;
Neodermata; Trematoda; Digenea; Echinostomida; Echinostomata;
Fascioloidea; Fasciolidae; Fasciola.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              -:- SIMILARITY: BELONGS TO THE GST SUPERFAMILY. ALPHA FAMILY.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      Query Match 15.7%; Score 681; DB 1; Length 217; Best Local Similarity 58.7%; Pred. No. 4e-30; Matches 125; Conservative 29; Mismatches 59; Indels
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             0099ElF59E49A49E CRC64;
                                                                                                                                       217 AA.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             Transferase, Antigen; Multigene family.
INIT_MET 0 BY SIMILARITY.
187 AIPQIDKYLKSSKYTAWPLQGWQATFGGGDHPP 219
                     PRT;
                                                                                                                                                                                                                                                                                                                                                                                                                                          MEDLINE=92155306; PubMed=1740183;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 MEDLINE=94039664; PubMed=8224094;
                                                                                                                                                                                                                                                                              Fasciola hepatica (Liver fluke)
                                                                                                                                                                               01-JUL-1993 (Rel. 26, Created)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               25196 MW;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                EMBL; M77680; AAA29139.1; -.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             EMBL; A00994; CAA00119.1; -. HSSP; P08515; 1GTB.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             SEQUENCE OF 8-217 FROM N.A.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         InterPro; IPR000521; GST. Pfam; PF00043; GST; 1.
                                                                                                                                         STANDARD;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         -! - SUBUNIT: HOMODIMER.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            196
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               217 AA;
                                                                                                                                                                                                                                                                                                                                                                                                                       SEQUENCE FROM N.A.
                                                                                                                                                                                                                                                                                                                                                                               NCBI_TaxID=6192;
                                                                                                                                                                                                                                                             CLASS-ALPHA).
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    Crameri S.;
                                                                                                                                         GT28_FASHE P31671;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           CONFLICT
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 SEQUENCE
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              ERRATUM.
                                                                                                                      GT28_FASHE
                                                                                                                                                            ò
```

0;

7 LGYWKIKGLVQPTRLLLEYLEEKYEEHLYERDEGDKWRNKKFELGLEFPNLPYYIDGDVK 66

δλ

```
This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation -
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            the Buropean Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See http://www.isb-sib.ch/announce/or send an email to license@isb-sib.ch).
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               Panaccio M., Wilson L.R., Crameri S.L., Wijffels G.L., Spithill T.W., "Molecular Characterization of cDNA sequences encoding glutathione S-transferases of Fasciola hepatica."; Exp. Parasitol. 74:232-237(1992).
                                       LTQSMAIIRYIADKHNMEGGCPKERAEISMLEGAVLDIRYGVSRIAYSKDFETLKVDFLS 126
                                                       127 KLPEMLKMFEDRLCHKTYLNGDHVTHPDFMLYDALDVVLYMDPMCLDAFPKLVCFKKRIE 186
                                                                                                                                        Panaccio M., Wilson L.R., Crameri S.L., Wijffels G.L., Spithill T.W., Exp. Parasitol. 77:385-385(1993).
4 LGYWKIRGLQQPVRLFLEYLGEEYEEHLYGRNDREKWLGDKFNMGLDLPNLPYYIDDKCK 63
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     Muro A., Rodriguez-Medina J.R., Hillyer G.V.; "Sequence analysis of a Fasciola hepatica glutathione S-transferase
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             Am. J. Trop. Med. Hyg. 48:457-463(1993).

-! FUNCTION: CONJUGATION OF REDUCED GLUTATHIONE TO A WIDE NUMBER OF EXOGENOUS AND ENDOGENOUS HYROPHOBIC ELECTROPHILES.

-! FUNCTION: GST ISOBLYTMES APPEAR TO PLAY A CENTRAL ROLE IN THE PARASITE DETOXIFICATION SYSTEM. OTHER FUNCTIONS ARE ALSO SUSPECTED INCLUDING A ROLE IN INCREASING THE SOLUBILITY OF HABMATIN IN THE PARASITE GUT.

-! CATALYTIC ACTIVITY: RX + GLUTATHIONE = HX + R-S-GLUTATHIONE.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    Eukaryota; Metazoa; Platyhelminthes; Turbellarian Platyhelminths;
                                                                                                                                                                                                                                                                                                                                                                            01-APR-1993 (Rel. 25, Created)
15-DEC-1998 (Rel. 37, Last sequence update)
15-DEC-1998 (Rel. 37, Last annotation update)
GLUTATHIONE S-TRANSFERASE 26 KDA 51 (EC 2.5.1.18) (GST51) (FH51)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    Rhabditophora; Eulecithophora; Revertospermata: Mediofusata;
Neodermata; Trematoda; Digenea; Echinostomida; Echinostomata;
Fascioloidea; Fasciolidae; Fasciola.
NCBL_TaxID=6192;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 -!- SIMILARITY: BELONGS TO THE GST SUPERFAMILY. ALPHA FAMILY.
                                                                                                                                                                                                                                                                                                                                         217 AA
                                                                                                                                                                                                                       187 AIPQIDKYLKSSKYIAWPLQGWQATFGGGDHPP 219
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       InterPro; IPR000521; GST.
Pfam; PF00043; GST; 1.
Transferase; Antigen; Multigene family.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             SUBCELLULAR LOCATION: CYTOPLASMIC
                                                                                                                                                                                                                                                                                                                                         PRT;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          MEDLINE=92155306; PubMed=1740183;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   MEDLINE=94039664; PubMed=8224094;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               SEQUENCE OF 9-217 FROM N.A. MEDLINE-93228188; PubMed-7682383;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 Fasciola hepatica (Liver fluke)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      EMBL; M77682; AAA29141.1; -.
                                                                                                                                                                                                                                                                                                                                         STANDARD;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            SUBUNIT: HOMODIMER.
                                                                                                                                                                                                                                                                                                                                                                                                                                                            (GST CLASS-ALPHA).
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      SEQUENCE FROM N.A.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             P08515;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                CDNA clone."
                                                                                                                                                                                                                                                                                                                                         GT26_FASHE
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  ERRATUM
                                                                                                                                                                                                                                                                                                                    GT26_FASHE
                                                                                                                                                                                                                                                                                                                                                             g
                                                                                                                                                       g
                                                                                                                                                                                                δ
                                                                                                                                                                                                                                   g
                                     ò
                                                                                                                 ò
```

```
MEDLINE-90094327; PubMed-2689439;
Townsend A.J., Goldsmith M.E., Pickett C.B., Cowan K.H.;
Trownsend A.J., Goldsmith M.E., Pickett C.B., Cowan K.H.;
"Isolation, characterization, and expression in Escherichia coli of
two murine Mu class glutathione S.transferase cDNAs homologous to the
rat subunits 3 (Ybi) and 4 (Yb2).";
J. Biol. Chem. 264:21582-21590(1989).
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            Pearson W.R., Reinhart J., Sisk S.C., Anderson K.S., Adler P.N.; "Tissue-specific induction of murine glutathione transferase mRNAs by butylated hydroxyanisole.";
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          67 LTQSMAIIRYIADKHNMLGGCPKERAEISMLEGAVLDIRYGVSRIAYSKDFETLKVDFLS 126
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            KLPEMLKMFEDRLCHKTYLNGDHVTHPDFMLYDALDVVLYMDPMCLDAFPKLVCFKKRIE 186
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               Gaps
                                                                                                                                                                                                                                                                                                                                                          7 LGYWKIKGLVQPTRLLLEYLEEKYEEHLYERDEGDKWRNKKFELGLEFPNLPYYIDGDVK 66
                                                                                                                                                                                                                                                                                                                                                                                            4 LGYWKIRGLQQPVRLLLEYLGEEYEEHLYGRDDREKWPGDKFNMGLDLPNLPYYIDDKCK 63
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            MEDLINE-83109018; PubMed-6822548; Pearson W.R., Windle J.J., Morrow J.F., Benson A.M., Talalay P.; Pearson W.R., Windle J.J., Morrow J.F., Benson A.M., Talalay P.; Increased synthesis of glutathione S-transferases in response to anticarcinogenic antioxidants. Cloning and measurement of messenger RNA.";
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              01-JUL-1989 (Rel. 11, Created)
01-JUL-1989 (Rel. 11, Last sequence update)
15-DEC-1998 (Rel. 37, Last annotation update)
GLUTATHIONE S-TRANSFERASE GT8.7 (EC 2.5.1.18) (GST 1-1) (GST CLASS-
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               Reinhart J., Pearson W.R.; "The structure of two murine class-mu glutathione transferase genes coordinately induced by butylated hydroxyanisole."; Arch. Biochem. Blophys. 303:383-393(1993).
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             Mus musculus (Mouse).
Bukaryota; McLazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Butheria; Rodentia; Sciurognathi; Muridae; Murinae: Mus.
                                                                                                                                                                                                                                                                                                 .,
                                                                                                                                                                                                                              15.6%; Score 680; DB 1; Length 217; 58.7%; Pred. No. 4.5e-30;
                                                                                                                                                                                                                                                                                                 56; Indels
BY SIMILARITY.

T -> S (IN REF. 3).

N -> D (IN REF. 3).

P -> T (IN REF. 3).

W: EFCA8FF7454EDC26 CRC64;
                                                                                                                                                                                                                                                                                                 Mismatches
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        217 AA
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      187 AIPQIDKYLKSSKYIAWPLQGWQATFGGGDHPP 219
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               Chem. 263:13324-13332(1988),
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        PRT;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 J. Biol. Chem. 258:2052-2062(1983)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    MEDLINE=93290350; PubMed=8512323;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             MEDLINE=88330838; PubMed=3417659;
                                                                                                                                                                                                                                                                                          32;
0
83
134
146
25242 MW;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 PRELIMINARY SEQUENCE OF 1-40.
                                                                                                                                                                                                                                                                                                 Conservative
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        STANDARD;
                                                                                                                                                                                                                                                         Best Local Similarity
Matches 125; Conserv
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               SEQUENCE FROM N.A.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      SEQUENCE FROM N.A.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   SEQUENCE FROM N.A.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               NCBI_TaxID=10090;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    GTM1_MOUSE
                              CONFLICT
CONFLICT
CONFLICT
SEQUENCE
                                                                                                                                                                                                                                 Query Match
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             Biol.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               P10649;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          127
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   MOUSE
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   GSTM1
   FT
FT
FT
SO
                                                                                                                                                                                                                                                                                                                                                             ŏ
                                                                                                                                                                                                                                                                                                                                                                                                                          qq
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       qq
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          δλ
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      ηq
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   ŏ
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                Dp
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               AC
DDT
DDT
DDT
DDE
DDE
DDE
DDE
DDE
DDE
RN DDE
RN N
RN DDE
RN N
RN DDE
RN N
RN DDE
RN N
RN DDE
DDE
RN DDE
RN DDE
RN DDE
RN DDE
RN DDE
RN DDE
RN DDE
RN DDE
RN DDE
RN DDE
RN DDE
RN DDE
RN DDE
RN DDE
RN DDE
RN DDE
RN DDE
RN DDE
RN DDE
RN DDE
RN DDE
RN DDE
RN DDE
RN DDE
RN DDE
RN DDE
RN DDE
RN DDE
RN DDE
RN DDE
RN DDE
RN DDE
RN DDE
RN DDE
RN DDE
RN DDE
DDE
RN DDE
RN DDE
RN DDE
RN DDE
RN DDE
RN DDE
RN DDE
RN DDE
RN DD
```

217 AA.

PRT;

```
Eukaryota; Métazoa; Chórdata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.
                                                                                01-AUG-1988 (Re). 08, Last sequence update)
15-JUL-1999 (Rel. 38, Last annotation update)
GLUTATHIONE S-TRANSFERASE YB1 (EC 2.5.1.18) (CHAIN 3) (GST M1-1)
                                                                                                                                                                                                                                                                   SEQUENCE FROM N.A. (CLONE PGTR200).
                                                                  13-AUG-1987 (Rel. 05, Created)
                                STANDARD;
                                                                                                                                                                             Rattus norvegicus (Rat).
                                                                                                                                                                                                                               NCBI_TaxID=10116;
                                                                                                                                         (GST CLASS-MU).
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      transferase 3-3
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               Pickett C.B.
                              GrM1_RAT
P04905;
               GTM1_RAT
RESULT
                                                                  This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See http://www.isb-sib.ch/announce/or send an email to license@isb-sib.ch).
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 ÷
                                                                "Identification of three classes of cytosolic glutathione transferase common to several mammalian species: correlation between structural
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          61 IDGDVKLTQSMAIIRYIADKHNMLGGCPKERAEISMLEGAVLDIRYGVSRIAYSKDFETL 120
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               5; Gaps
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       6 ILGYWKIKGLVQPTRLLLEYLEEKYEEHLYERDEG-----DKWRNKKFELGLEFPNLPYY 60
                                                                                                                                                                      STRAINCD-1; TISSUE-Liver;
MEDLINE=96189427; PubMed=8605288;
MEDLINE=96189427; PubMed=8605288;
Mitchell A. E., Worin D., Lame M. W., Jones A.D.;
Purification of murine glutathione S-transferases.";
Chem. Res. Toxicol. 8:1054-1062(1995).
-!- FUNCTION: CONJUGATION OF REDUCED GLUTATHIONE TO A WIDE NUMBER OF EXOCENOUS AND ENDOGRNOUS HYDROPHOBIC ELECTROPHILES.
-!- CATALYTIC ACTIVITY: RX + GLUTATHIONE = HX + R-S-GLUTATHIONE.
                                                                                                                                                                                                                                                                                                                                                    SUBCELLULAR LOCATION: CYTOPLASMIC.
MASS SPECTROMETY: MW-25838.4; MW-ERR-2; METHOD-ELECTROSPRAY.
SIMILARITY: BELONGS TO THE GST SUPERFAMILY. MU FAMILY.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         11.1%; Score 481.5; DB 1; Length 217; 44.4%; Pred. No. 1.9e-19; Live 38; Mismatches 71; Indels 5;
         MEDLINE-86042634; PubMed-3864155;
Mannervik B., Alin P., Guthenberg C., Jensson H., Tahir M.K.,
Warholm M., Joernvall H.;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       217 AA; 25839 MW; AIEE3938F590B829 CRC64;
                                                                                                     data and enzymatic properties.";
Proc. Natl. Acad. Sci. U.S.A. 82:7202-7206(1985).
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                L13448; -; NOT_ANNOTATED_CDS.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           PIR; A20831; A20831.
PIR; A28946. A28946.
PIR; 124735; I34735.
PIR; 333860. S33860.
HSSP; P04905; 6GSY.
SWISS-2DPAGE; P10649; MOUSE.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             Pfam; PF00043; GST; 1.
PRINTS; PR01267; GSTRNSFRASEM.
Transferase; Multigene family.
PRELIMINARY SEQUENCE OF 1-24.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          EMBL; J03952; AAA37747.1; -. EMBL; J04632; AAA37705.1; -.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          InterPro; IPR000521; GST.
InterPro; IPR003081; GST_mu.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               91; Conservative
                                                                                                                                                                                                                                                                                                                                          SUBUNIT: HOMODIMER.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           MGD; MGI:95860; Gstml
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           Query Match
Best Local Similarity
Matches 91; Conserva
                                                                                                                                                          CHARACTERIZATION.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       SEQUENCE
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        INIT_MET
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              EMBL;
```

```
"Rat liver glutathione S-transferases. DNA sequence analysis of a Yb2 cDNA clone and regulation of the Yb1 and Yb2 mRNAs by phenobarbital.";
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                Hisieh J.-C., Huang S.-C., Chen W.-L., Lai Y.-C., Tam M.F.;
"Cysteine 86 is not needed for the enzymic activity of glutathione S-
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    "I X. Zhang P., Armstrong R.N., Gilliland G.L.;
"The three-dimensional structure of a glutathione S-transferase from the mu gene class. Structural analysis of the binary complex of isoenzyme 3-3 and qlutathione at 2.2-A resolution.";
Biochemistry 31:10169-10184(1992).
                                     Lai H.-C.J., Grove G., Tu C.-P.D.; "Cloning and sequence analysis of a cDNA for a rat liver glutathione
                                                                                                                                                                                                                                                                MEDLINE=86033768; PubMed=3840477;
Ding G.J.-F., Lu A.Y.H., Pickett C.B.;
"Rat liver glutathione S-transferases. Nucleotide sequence analysis of a Ybl cDNA clone and prediction of the complete amino acid sequence of the Ybl subunit.";
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       X-RAY CRYSTALLOGRAPHY (2.5 ANGSTROMS).
Fu J.-H., Rose J., Tam M.F., Wang B.-C.;
New crystal forms of a mu-class glutathione S-transferase from rat
liver."
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              purification, molecular properties and odorant biotransformation."; Biochem. J. 292:379-384(1993).
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                Ding G.J.-F., Ding V.D.-H., Rodkey J.A., Bennett C.D., Lu A.Y.H.,
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    Chang C., Saltzman A.G., Sorensen N.S., Hilpakka R.A., Liao S.; "Identification of glutathione S-transferase Yb1 mRNA as the androgen-repressed mRNA by CDNA cloning and sequence analysis."; J. Biol. Chem. 262:11901-11903(1987).
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               STRAIN=WISTAR; TISSUE=Olfactory epithelium;
MEDLLNE=92277499; PubMed=9503877;
Hen-Arie N., Khen M., Lancet D.;
"Glutathione S-transferases in rat olfactory epithelium;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 X-RAY CRYSTALLOGRAPHY (2.2 ANGSTROMS).
                                                                                                                                                  Nucleic Acids Res. 14:6101-6114(1986).
                                                                                                                                                                                                                                                                                                                                                                                                                                                                     Biol. Chem. 260:13268-13271(1985).
                                                                                                                                                                                                                                      (CLONE PGTA/C44).
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             Biol. Chem. 261:7952-7957(1986).
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          MEDLINE=86224097; PubMed=3011803;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          MEDLINE=91354218; PubMed=1883338;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           MEDLINE=93041702; PubMed=1420139;
MEDLINE=86312882; PubMed=2875437;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   MEDLINE=87308179; PubMed=3040722;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     Biochem, J. 278:293-297(1991).
                                                                                                                     S-transferase Yb subunit.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        MUTAGENESIS OF CYS-86
                                                                                                                                                                                                                                  SEQUENCE FROM N.A.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                SEQUENCE FROM N.A.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             SEQUENCE FROM N.A.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         SEQUENCE OF 1-23.
```

g δ g ò Dp δλ

```
q
                                                                                                                                                                                                                                                                                                                                                                                Dp
                                                                                                                                                                                                                                                                                                               δŏ
                                                                                                                                                                                                                                                                                                                                    qq
                                                                                                                                                                                                                                                                                                                                                           ŏ
                                                                                                                                                                                                                                                                                                                                                                                                      ŏ
                                                                                                                                                                                                                                                               This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation - the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license girement (See http://www.isb-sib.ch/announce/or send an email to license@isb-sib.ch).
                          MEDLINE-94153886; PubMed=8110735;
M. MEDLINE-94153886; PubMed=8110735;
Ji X., Johnson W. W., Sesay M.A., Dickert L., Prasad S.M., Ammon H.L.,
Armstrong R.W., Gilliland G.L.;
Structure and function of the xenobiotic substrate binding site of a glutathione S-transferase as revealed by X-ray crystallographic analysis of product complexes with the diastereomers of 9-(5-
The glutathionyl)-10-hydroxy-9.10-dihydrophenanthrene.", MIDE NUMBER
C. -I- FUNCTION: CONJUGATION OF REDUCED GLUTATHIONE TO A WIDE NUMBER
OF EXORENOUS AND ENDOGENOUS HYDROPHOBIC ELECTROPHILES.
-I- FUNCTION: THE OLFACTORY GST MAY BE CRUCIAL FOR THE ACUITY OF THE OLFACTORY PROCESS.
-I- FUNCTION: THE OLFACTORY GST MAY BE CRUCIAL FOR THE ACUITY OF THE CATALYTIC ACTIVITY: RX + GLUTATHIONE = HX + R-S-GLUTATHIONE.
-I- SUBGRILLANE MOCATION: CYTOPLASMIC.
-I- SUBGRILLANE MOCATION: CYTOPLASMIC.
-I- SUBGRILLANEOUS: RAT LIVER GLUTATHIONE S-TRANSFERASES (GST) CONSIST OF BIRNRY COMBINATIONS OF THREE MAJOR CLASSES OF SUBUNITS DESIGNATED: YA (25 KDA), YB (27 KDA) AND YC (28 KDA).
-I- SIMILARITY: BELONGS TO THE GST SUBERFAMILY. MU FAMILY.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   C->S: NO CHANGE IN ACTIVITY.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               PRINTS; PR01267; GSTRNSFRASEM.
Transferase; Multigene family; 3D-structure; Olfaction.
INIT_MET 0 0
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            I -> N (IN REF. 3).
KS -> NC (IN REF. 2).
                      X-RAY CRYSTALLOGRAPHY (2.0 ANGSTROMS)
Acta Crystallogr. D 50:219-224(1994).
                                                                                                                                                                                                                                                                                                                                                       EMBL; X04229; CAA27811.1; -. EMBL; M11719; AAA41287.1; -. EMBL; J02810; AAA41293.1; -.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               IPRO00521; GST.
IPR003081; GST_mu.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            168
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             113
222
224
332
446
550
550
70
70
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          08-NOV-96.
27-JAN-99.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       PF00043; GST; 1.
                                                                                                                                                                                                                                                                                                                                                                                                                                                       31-OCT-93.
31-JAN-94.
31-OCT-93.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    08-NOV-96.
08-NOV-96.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    08-NOV-96.
08-NOV-96.
                                                                                                                                                                                                                                                                                                                                                                                                                                      31-0CT-93
                                                                                                                                                                                                                                                                                                                                                                                                                                               31-0CT-93
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           31-OCT-93
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          96-NON-80
                                                                                                                                                                                                                                                                                                                                                                                                   A25510; A25510.
A29794; A29794.
                                                                                                                                                                                                                                                                                                                                                                                          A24085; A24085
                                                                                                                                                                                                                                                                                                                                                                                                                          S17167
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            168
198
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            111
123
123
137
137
147
167
167
167
                                                                                                                                                                                                                                                                                                                                                                                                                          S17167;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    6GST;
6GSU;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          6GSV;
6GSW;
6GSX;
                                                                                                                                                                                                                                                                                                                                                                                                                                      GSB;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                               GST;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                InterPro;
InterPro;
                                                                                                                                                                                                                                                                                                                                                                                                                                                GSC;
                                                                                                                                                                                                                                                                                                                                                                                                                                                            GST;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                      GST;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            GST;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           egsy;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      5FWG;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      INIT_MET
MUTAGEN
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           CONFLICT
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             STRAND
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              STRAND
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        STRAND
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  STRAND
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      HELIX
TURN
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  HELIX
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       Pfam;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         TURN
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     TURN
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               TURN
                                                                                                                                                                                                                                                                                                                                                                                                                PIR;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                    PDB;
PDB;
PDB;
PDB;
PDB;
PDB;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             FURN
                                                                                                                                                                                                                                                                                                                                                                                                                                      PDB;
                                                                                                                                                                                                                                                                                                                                                                                                                                                PDB;
```

```
1;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation -
                                                                                                                                                                                                                                                                                                                                                                                61 IDGDVKLTQSMAIIRYTADKHNMLGGCPKERAEISMLEGAVLDIRYGVSRIAYSKDFETL 120
                                                                                                                                                                                                                                                                                                                                                                                                121 KVDFLSKLPEMLKMFEDRLCHKTYLNGDHVTHPDFMLYDALDVVLYMDPMCLDAFPKLVC 180
                                                                                                                                                                                                                                                                                                                                                                                                                                                     Gaps
                                                                                                                                                                                                                                                                                                                                           6 ILGYWKIKGLVQPTRLLLEYLEEKYEEHLYERDEG-----DKWRNKKFELGLEFPNLPYY 60
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             01-NOV-1995 (Rel. 32, Last sequence update)
15-JUL-1999 (Rel. 38, Last annotation update)
GLUTATHIONE S-TRANSFERASE MU 1 (EC 2.5.1.18) (GST MU I) (GST CLASS-
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         -:- CATALYTIC ACTIVITY: RX + GLUTATHIONE = HX + R-S-GLUTATHIONE. -:- SUBUNIT: HOMODIARR. -:- SUBCELLUIAR LOCATION: CYTOPLASMIC. -:- SUBCELLUIAR LOCATION: CYTOPLASMIC. -:- TISSUE SPECIFICITY: WELL EXPRESSED IN RABBIT LIVER, BRAIN, AND
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             Oryctolagus cuniculus (Rabbit).
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Lagomorpha; Leporidae; Oryctolagus.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    S-transferase from rabbit liver.";
Arch. Biochem. Biophys. 318 424 4 429 (1995).
-!- UNCTION: CONJUGATION OF REDUCED GLUTATHIONE TO A WIDE NUMBER OF EXOGENOUS AND ENDOGENOUS HYDROPHOBIC ELECTROPHILES.
                                                                                                                                                                                                                                                                                                  5;
                                                                                                                                                                                                                                                                        11.1%; Score 481.5; DB 1; Length 217; 44.4%; Pred. No. 1.9e-19;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            MEDLINE=95251394; PubMed=7733673;
Lee S.H., Lee S.H., Han J.S., Kim Y.S., Koh J.K.;
"Cloning and expression of a cDNA for mu-class glutathione
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             -!- SIMILARITY: BELONGS TO THE GST SUPERFAMILY. MU FAMILY.
                                                                                                                                                                                                                                                                                                  Indels
                                                                                                                                                                                                                                25782 MW; 2ACE8D49DA785118 CRC64;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       217 AA
                                                                                                                                                                                                                                                                                                 38; Mismatches
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       PRT;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            FKKRIEAIPQIDKYLKSSKYIAWPL 205
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          183 FLARFEGLKKISAYMKSSRYLSTPI 207
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              01-NOV-1995 (Rel. 32, Created)
01-NOV-1995 (Rel. 32, Last sequ
15-JUL-1999 (Rel. 38, Last anno
                                                                                                                                                                                                                                                                                      Best Local Similarity 44.4%;
                                                                                                                                                                                                                                                                                                  91; Conservative
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       STANDARD;
217 AA;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   SEQUENCE FROM N.A.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           NCBI_TaxID=9986;
TISSUE=Liver;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     GTMU_RABIT
P46409;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                KIDNEY
                                                                                                                                                                                                                                  SEQUENCE
                                                                                                                                                                                                                                                                         Query Match
                                                  HELIX
TURN
                                                                                                        TURN
HELIX
TURN
                                                                                                                                                HELIX
TURN
                                                                             HELIX
TURN
            HELIX
TURN
                                                                                                                                                                          HELIX
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            RESULT 10
                                                                                                                                                                                                     TURN
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           GTMU_RABIT
                                                                                                                                                                                          TURN
                                                                                                                                                                                                                                                                                                  Matches
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            181
                                        TURN
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        \widehat{\mathbb{R}}
```

us-09-724-296-2.rsp

8

```
qq
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       òγ
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        q
                use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See http://www.isb-sib.ch/announce/or send an email to license@isb-sib.ch).
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                ;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   MEDLINE-93228631; PubMed-8471052; Zhong S., Spurr N.K., Hayes J.D., Wolf C.R.; Thong S., Spurr N.K., Hayes J.D., Wolf C.R.; Meduced amino acid sequence, gene structure and chromosomal location of a novel human class Mu glutathione S-transferase, GSTM4."; Biochem. J. 291:41-50(1993).
There are no restrictions on
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               64 DGTHKLTQSNAILRYLARKHGLCGETEEERIRVDILENQLMDNRFQLVNVCYSPDFEKLK 123
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        122 VDFLSKLPEMLKMFEDRLCHKTYLNGDHVTHPDFMLYDALDVVLYMDPMCLDAFPKLVCF 181
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    62 DGDVKLTQSMAIIRYIADKHNMLGGCPKERAEISMLEGAVLDIRYGVSRIAYSKDFETLK 121
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        124 PEYLKGLPEKLOLYSQFLGSLPWFAGDKITFADFLVYDVLDQNRIFVPGCLDAFPNLKDF 183
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             7 LGYWKIKGLVQPTRLLLEYLEEKYEEHLYERDEG-----DKWRNKKFELGLEFPNLPYYI 61
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              4 LGYWDVRGLALPIRMLLEYTDTSYEEKKYTMGDAPNYDQSKWLSEKFTLGLDFPNLFPYLI 63
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    MEDLINE=93352467; PubMed=8349586; Comstock K.E., Johnson K.J., Rifenbery D., Henner W.D.; Comstock K.E., Johnson K.J., Rifenbery D., Henner W.D.; Isolation and analysis of the gene and cDNA for a human Mu class glutathione S-rransferase, GSTM4 "; J. Biol. Chem. 268:16958-16965(1993).
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    01-NOV-1995 (Rel. 32, Last sequence update)
20-AUG-2001 (Rel. 40, Last annotation update)
GLUTATHIONE S-TRANSFERASE MU 4 (EC 2.5.1.18) (GSTM4-4) (GTS-MU2)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 Taylor J.B., Oliver J., Sherrington R., Pemble S.E.; "Structure of human glutathione S-transferase class Mu genes."; Biochem, J. 274:587-593(1991).
                                                                                                                                                                                                                                                                                                                                                                                                                                          11.0%; Score 477.5; DB 1; Length 217; 44.6%; Pred. No. 3e-19;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             77; Indels
                                                                                                                                                                                                                                                                                                                                                                 785EA932C543ECB2 CRC64;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             36; Mismatches
                                                                                                                                                                                                                                                                                                                                      BY SIMILARITY
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       184 HVRFEGLPKISAYMKSSRFIRVPVFLKKATWTG 216
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            182 KKRIEAIPQIDKYLKSSKYIAWPLQGWQATFGG 214
  European Bioinformatics Institute.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           TISSUE-Lymphocytes;
MEDLINE-91174774; PubMed-2006920;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           01-NOV-1995 (Rel. 32, Created)
01-NOV-1995 (Rel. 32, Last sequ
                                                                                                                                                                                                                                                                                PRINTS; PR01267; GSTRNSFRASEM
Transferase; Multigene family
                                                                                                                                                                                                                                                                                                                                                                 SEOUENCE 217 AA; 25286 MW;
                                                                                                                                                                                                    InterPro; IPR000521; GST.
InterPro; IPR003081; GST_mu.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   SEQUENCE OF 39-160 FROM N.A.
                                                                                                                                                    EMBL; L23766; AAA69665.1; -.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                     Best_Local Similarity 44.69
Matches 95; Conservative
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              STANDARD;
                                                                                                                                                                                                                                                          Pfam; PF00043; GST; 1.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        Homo sapiens (Human).
                                                                                                                                                                               1HNB.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       SEQUENCE FROM N.A.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   SEQUENCE FROM N.A.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      NCBI_TaxID=9606;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       CLASS-MU).
                                                                                                                                                                                  HSSP; P28161
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              GTM4_HUMAN
Q03013;
                                                                                                                                                                                                                                                                                                                                                                                                                                          Query Match
                                                                                                                                                                                                                                                                                                                                      INIT_MET
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                GSTM4.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     GTM4_HUMAN
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       (GST
  qq
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    ŏλ
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        Qγ
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      Dp
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            δ
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            q
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                HERE REPAREMENT OF THE REPAREM
```

```
This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See http://www.isb-sib.ch/announce/or send an email to license@isb-sib.ch).
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            01-APR-1993 (Rel. 25, Created)
01-APR-1993 (Rel. 25, Last sequence update)
20-AUG-2001 (Rel. 40, Last annotation update)
GLUTATHIONE S-TRANSFERASE MU 2 (EC 2.5.1.18) (GSTM2-2) (GST CLASS-MU).
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  62 DGDVKLTQSMAIIRYIADKHNMLGGCPKERAEISMLEGAVLDIRYGVSRIAYSKDFETLK 121
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            122 VDFLSKLPEMLKMFEDRLCHKTYLNGDHVTHPDFMLYDALDVVLYMDPMCLDAFPKLVCF 181
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          Gaps
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  7 LGYWKIKGLVQPTRLLLEYLEEKYEEHLYERDEG-----DKWRNKKFELGLEFPNLPYYI 61
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          5 LGYWDIRGLAHAIRLLLEYTDSSYEEKKYTMGDAPDYDRSQWLNEKFKLGLDFPNLPYLI
              Comstock K.E., Widersten M., Hao X.Y., Henner W.D., Mannervik B.; "A comparison of the enzymatic and physicochemical properties of human glutathione transferase M4-4 and three other human Mu class enzymes.";
                                                                                                                                             Arch. Biochem. Biophys. 311:487-495(1994).
-!- FUNCTION: CONJUGATION OF REDUCED GLUTATHIONE TO A WIDE NUMBER OF EXOGENOUS AND ENDOGENOUS HYDROPHOBIC ELECTROPHILES.
                                                                                                                                                                                                                                                           -!- CATALYTIC ACTIVITY: RX + GLUTATHIONE = HX + R-S-GLUTATHIONE
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            .
2
                                                                                                                                                                                                                                                                                                                                                -!- TISSUE SPECIFICITY: EXPRESSED IN A WIDE VARIETY OF TISSUES.
-!- SIMILARITY: BELONGS TO THE GST SUPERFAMILY. MU FAMILY.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    Length 218;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           SMT -> PM1 (IN REF. 2).

I -> M (IN REF. 1; AAA57346).

D -> G (IN REF. 2).

C -> M (IN REF. 2).

35E96FA54D566BIE CRC64;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         77; Indels
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              DB 1;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 10.9%; Score 474.5; DB 1 43.9%; Pred. No. 4.4e-19;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             217 AA
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            37; Mismatches
                                                                                                                                                                                                                                   ACTIVE ON 1-CHLORO-2,4-DINITROBENZENE.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             182 KKRIEAIPQIDKYLKSSKYIAWPLQGWQATFG 213
                                                                                                                                                                                                                                                                                                                      -!- SUBCELLULAR LOCATION: CYTOPLASMIC.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             PRT;
MEDLINE=94263230; PubMed=8203914;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   17 17 1
37 37 1
212 212 V
218 AA; 25561 MW;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         PRINTS; PR01257; GSTRNSFRASEM.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    Transferase; Multigene family
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           EMBL; M96234; AAA57347.1; -. EMBL; M96233; AAA57346.1; -. EMBL; X68677; CAA48637.1; -. EMBL; X56837; CAA40167.1; -.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    GST_mu.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            Best Local Similarity 43.9
Matches 93; Conservative
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             STANDARD;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      nterPro; IPR000521; GSF
                                                                                                                                                                                                                                                                                              !- SUBUNIT: HOMODIMER.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    nterPro; IPR003081;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              PF00043; GST;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 HSSP; P28161; 1HNB.
MIM; 138333; -.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               GSTM2 OR GST4.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             GTM2_HUMAN
P28161;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 CONFLICT
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         CONFLICT
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 SEQUENCE
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       Query Match
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    Pfam;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               GTM2_HUMAN
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       9
```

σ

÷

```
qq
                                                                                                                                                                      ò
                                                                                                                                                                                     qq
                                                                                                                                                                                                     οy
                                                                                                                                                                                                                     QQ
                                                                                                                                                                                                                                     δ
                                                                                                                                                                                                                                                    qq
FT
FT
FT
FT
FT
FT
SQ
                                                                                                                                                                                                                     This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way
                                                                                                                                                                                                                                           use by non-profit institutions as long as its content is in o way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See http://www.isb-sib.ch/announce/or send an email to license@lsb-sib.ch).
                                                                                                             Reghunathan S., Chandross R.J., Kretsinger R.H., Allison T.J., Penington C.J., Rule G.S., "Crystal structure of human class mu glutathione transferase GSTM2-2.
                                                     Vorachek W.R., Pearson W.R., Rule G.S.;
"Cloning, expression, and characterization of a class-mu glutathione transferase from human muscle, the product of the GST4 locus.";
Proc. Natl. Acad. Sci. U.S.A. 88:4443-4447(1991).
                                                                                                                                     Bukaryota, Metazoa, Chordata, Craniata, Vertebrata, Euteleostomi,
Mammalia, Eutheria, Primates, Catarrhini, Hominidae, Homo.
                                                                                                                                                                                     -:- SUBCELLUAR LOCATION: CYTOPLASMIC.
-:- TISSUE SPECIFICITY: THIS IS A MUSCLE ISOZYME.
-:- SIMILARITY: BELONGS TO THE GST SUPERFAMILY. MU FAMILY.
                                                                                                                                                                                                                                                                                                                                                                      PRINTS; PR01267; GSTRNSFRASEM.
Transferase; Multigene family; 3D-structure.
                                                                                                X-RAY CRYSTALLOGRAPHY (1.85 ANGSTROMS).
                                                 MEDLINE-91239584; PubMed-2034681;
                                                                                                        MEDLINE=94238693; PubMed=8182750;
                                                                                                                                                                                                                                                                                                                                                       IPR003081; GST_mu.
                                                                                                                                                                                                                                                                                       EMBL; M63509; AAA60963.1; -.
                                                                                                                                                                                                                                                                                                                                               InterPro; IPR000521; GST
                                                                                                                                                                                                                                                                                                                                                                                                               1HNA; 31-JAN-94.
1HNB; 31-JAN-94.
1HNC; 31-JAN-94.
2GTU; 02-MAR-99.
                                                                                                                                                                                                                                                                                                                                                                GST; 1.
  sapiens (Human).
                                                                                                                                                                                                                                                                                                 PIR; A39375; A39375.
                                         SEQUENCE FROM N.A.
                                                                                                                                                                                                                                                                                                                                                                                                               111
113
223
227
227
328
338
441
441
441
72
65
65
67
67
72
90
90
                         NCBI_TaxID=9606;
                                                                                                                                                                                                                                                                                                                                                                PF00043;
                                                                                                                                                                                                                                                                                                                                        138380;
                                                                                                                                                                                                                                                                                                                                                         InterPro;
                                                                                                                                                                                                                                                                                                                                                                                        INIT_MET
STRAND
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              STRAND
HELIX
TURN
                                                                                                                                                                                                                                                                                                                                                                                                                                                                        STRAND
HELIX
                                                                                                                                                                                                                                                                                                                                                                                                                                               STRAND
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 STRAND
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 STRAND
                                                                                                                                                                                                                                                                                                                                                                                                         STRAND
                                                                                                                                                                                                                                                                                                                                                                                                                                         STRAND
                                                                                                                                                                                                                                                                                                                                                                                                                        HELIX
TURN
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         HELIX
TURN
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 HELIX
TURN
                                                                                                                                                                                                                                                                                                                                                                                                                rurn
                                                                                                                                                                                                                                                                                                                                                                                                                                                                 TURN
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        TURN
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        TURN
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          rurn
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          rurn
                                                                                                                                                                                                                                                                                                                                 PDB;
                                                                                                                                                                                                                                                                                                         PDB;
                                                                                                                                                                                                                                                                                                                         PDB;
```

```
64 DGTHKITQSNAILRYIARKHNLCGESEKEQIREDILENQFMDSRMQLAKLCYDPDFEKLK 123
                                                                                                                                                                                                                                                                                                                                                                                                                                               122 VDFLSKLPEMLKMFEDRLCHKTYLNGDHVTHPDFMLYDALDVVLYMDPMCLDAFPKLVCF 181
                                                                                                                                                                                                                                                                                                                                                                                                                                                                  DGDVKLTQSMAIIRYIADKHNMLGGCPKERAEISMLEGAVLDIRYGVSRIAYSKDFETLK 121
                                                                                                                                                                                                                                                                   Gaps
                                                                                                                                                                                                                                                                                                                               4 LGYWNIRGLAHSIRLLLEYTDSSYEEKKYTMGDAPDYDRSQWLNEKFKLGLDFPNLPYLI 63
                                                                                                                                                                                                                                                                                                   7 LGYWKIKGLVQPTRLLLEYLEEKYEEHLYERDEG-----DKWRNKKFELGLEFPNLPYYI 61
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   Kamei K., Oshino R., Hara S.;
"Amino acid sequence of glutathione S-transferase b from guinea pig
iiver ".
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Rodentia; Hystricognathi; Caviidae; Cavia
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            J. Biochem. 107:111-117(1990).
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      01-AUG-1990 (Rel. 15, Created)
01-AUG-1990 (Rel. 15, Last sequence update)
01-NOV-1995 (Rel. 32, Last annotation update)
GLUTATHIONE S-TRANSFERASE B (EC 2.5.1.18) (GST CLASS-MU).
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  OF EXOGENOUS AND ENDOGENOUS HYDROPHOBIC ELECTROPHILES.
                                                                                                                                                                                                                                                                   5,
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          Lower 4/3.5; DB 1; Length 217;
Pred. No. 5e-19;
                                                                                                                                                                                                                           Length 217;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              -:- SUBUNIT: HOMODIMER.
-:- SUBCELLULAR LOCATION: CYTOPLASMIC.
-:- SIMILARITY: BELONGS TO THE GST SUPERFAMILY. MU FAMILY.
                                                                                                                                                                                                                                                                   Indels
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                PERNYS; PRO1267; GSTRNSFRASEM.
Transferase; Multigene family.
SEQUENCE 217 AA; 25719 MW; D29F7951D4E9365E CRC64;
                                                                                                                                                                            25613 MW; 25603A909482CA39 CRC64;
                                                                                                                                                                                                                             10.9%; Score 473.5; DB 1;
44.3%; Pred. No. 5e-19;
iive 35; Mismatches 78;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    217 AA
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         182 KKRIEAIPQIDKYLKSSKYIAWPLOGWQATFG 213
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         184 ISRFEGLEKISAYMKSSRFLPRPVFTKMAVWG 215
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      PRT;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        MEDLINE=90236961; PubMed=2332413;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               Cavia porcellus (Guinea pig).
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             10.9%;
45.1%;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       PIR, JX0095, JX0095.
HSSP, P04905, 6GSY.
Interpro; IPR000521, GST.
Interpro; IPR003081; GST_mu.
                                                                                                                                                                                                                                                                     Conservative
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      STANDARD;
147
1149
1150
1150
1169
1176
1197
1197
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   PF00043; GST;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                Best Local Similarity
                                                                                                                                                                               217 AA;
                                                                                                                                                                                                                                                     Similarity
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  NCBI_TaxID=10141;
                 148
150
151
151
171
171
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           TISSUE=Liver;
                                                                                                                                                                                                                                                                     94;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    GTMU_CAVPO
P16413;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                Query Match
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        SEQUENCE
                                                                                                                                                                                                                                   Query Match
                                                                                                                                                                                 SEQUENCE
                                                                                                                                                                                                                                                       Best Local
                                   TURN
STRAND
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   liver.
                    STRAND
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      Pfam;
                                                                       HELIX
HELIX
HELIX
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      GTMU_CAVPO
                                                                                                                            HELIX
                                                                                                                                                                                                                                                                        Matches
                                                                                                                                             TURN
                                                                                                                                                                                                                                                                                                                                                                               62
   TURN
                                                                                                                                                                FURN
```

```
This SWISS-PROT entry is copyright. It is produced through a collaboration
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   between the Swiss Institute of Bioinformatics and the EMBL outstation—the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See http://www.isb-sib.ch/announced.
                                            61 IDGDVKLTQSMAIIRYIADKHNMLGGCPKERAEISMLEGAVLDIRYGVSRIAYSKDFETL 120
                                                                 121 KVDFLSKLPEMLKMFEDRLCHKTYLNGDHVTHPDFMLYDALDVVLYMDPMCLDAFPKLVC 180
                                                                                                                                                        123 KPEFLKTIPEKMKMYSEFLGKRPWFAGDKVTLCGFLAYDVLDQYQMFEPKCLDPFNLKD 182
           3 ILGYWNVRGLTNPIRLLLEYTDSSYEEKKYTMGDAPDSDRSQWLNEKFKLGLDFPNLPYL 62
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 "Cloning of a mu-class glutathione S-transferase complementary DNA and characterization of its glucocorticoid inducibility in a smooth
                                                                                                                                                                                                                                                                                                                                                                                                                                                  Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Cricetinae;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              MEDLINE=92335246; PubMed=1631097;
Fan W.M., Trifiletti R., Norris J.S., Cooper T.M.;
"Cloning of a mu-class glutathione S-transferase gene and
identification of the glucocorticoid regulatory domains in its 5'
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             Norris J.S., Schwartz D.A., Macleod S.L., Fan W.M., O'Brien T.J., Harris S.E., Trifiletti R., Cornett L.E., Cooper T.M., Levi W.M.,
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          -!- FUNCTION: CONJUGATION OF REDUCED GLUTATHIONE TO A WIDE NUMBER OF EXOGENOUS AND ENDOGENOUS HYDROPHOBIC ELECTROPHILES.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              -!- CATALYTIC ACTIVITY: RX + GLUTATHIONE = HX + R-S-GLUTATHIONE.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           -!- SUBCELLULAR LOCATION: CYTOPLASMIC.
                                                                                                                                                                                                                                                                                                                                                                                                              GLUTATHIONE S-TRANSFERASE (BC 2.5.1.18) (GST CLASS-MU). Mesocricetus auratus (Golden hamster).
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      Proc. Natl. Acad. Sci. U.S.A. 89:6104-6108(1992).
                                                                                                                                                                                                                                                                                                                                                                             (Rel. 25, Last sequence update)
                                                                                                                                                                                                                                                                                                                           217 AA.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               or send an email to license@isb-sib.ch).
                                                                                                                                                                                                                                                                                                                             PRT;
                                                                                                                                                                                                               181 FKKRIEAIPQIDKYLKSSKYIAWPL 205
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             MEDLINE=92049380; PubMed=1944302;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       muscle tumor cell line.";
Mol. Endocrinol. 5:979-986(1991).
                                                                                                                                                                                                                                                                                                                                                          01-APR-1993 (Rel. 25, Created)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                PRINTS; PR01267; GSTRNSFRASEM.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             Transferase; Multigene family.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  EMBL; M59772; AAA37075.1; -.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     EMBL; X61033; CAA43368.1; -.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           InterPro; IPR003081; GST_mu.
                                                                                                                                                                                                                                                                                                                           STANDARD;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          InterPro; IPR000521; GST
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                SUBUNIT: HOMODIMER.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                TlSSUE=Smooth muscle;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              TISSUE=Smooth muscle;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           PF00043; GST; 1
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  PIR; A23732; A23732.
PIR; S21908; S21908.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       flanking sequence."
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         HSSP; PO4905; 6GST.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             SEQUENCE FROM N.A.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         SEQUENCE FROM N.A.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          NCBI_TaxID=10036;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         Mesocricetus.
                                                                                                                                                                                                                                                                                                                           CTMU_MESAU
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   Smith R.G.;
                                                                                                                                                                                                                                                                                                        GTMU MESAU
                                                                                Ωþ
                                                                                                                                                                                                                                                                                                                                           ŏ
                                                                                                                                                                                            δ
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation - the Buropean Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See http://www.isb-sib.ch/announce/or send an email to license@isb-sib.ch).
                                                                                                                                                                                                          122 VDFLSKLPEMLKMFEDRLCHKTYLNGDHVTHPDFMLYDALDVVLYMDPMCLDAFPKLVCF 181
                                                                                                       62 DGDVKLTQSMAIIRYIADKHNMLGGCPKERAEISMLEGAVLDIRYGVSRIAYSKDFETLK 121
                                                                   4 LGYWNIRGLTHPIRLILEYTNSGYEEKRYNMGDAPDYDRSQWLNEKFKLGLDFPNLPYLI 63
       5; Gaps
                                7 LGYWKIKGLVQPTRLLLEYLEEKYEEHLYERDEG-----DKWRNKKFELGLEFPNLPYYI 61
                                                                                                                                                                                                                                                                                                                                                                                                                  01-JUL-1993 (Rel. 26, Created)
01-JUL-1993 (Rel. 26, Last sequence update)
15-DEC-1998 (Rel. 37, Last annotation update)
16-DEC-1998 (Rel. 37, Last annotation update)
CLUTAPHIONE S-TRANSFERASE VI (EC 2.5.1.18) (CHAIN 3) (GST CLASS-MU).
Cricetulus longicaudatus (Long-tailed hamster) (Chinese hamster)
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Cricetinae;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         Gaps
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 6 ILGYWKIKGLVQPTRLLLEYLEEKYEEHLYERDEG-----DKWRNKKFELGLEFPNLPYY 60
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            -i - FUNCTION: CONJUGATION OF REDUCED GLUTATHIONE TO A WIDE NUMBER OF EXOGENOUS AND ENDOGENOUS HYDROPHOBIC ELECTROPHILES.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    MEDLINE=91031445; PubMed=2226437; de Saint Vincent B.R., Hyrien O., Debatisse M., Buttin G.; "Coamplification of mu class glutathione S-transferase genes and adenylate deaminase gene in coformycin-resistant Chinese hamster
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    -!- CATALYTIC ACTIVITY: RX + GLUTATHIONE = HX + R-S-GLUTATHIONE.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             5;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   DB 1; Length 217;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     -!- SÜBCELLULAR LOCATION: CYTOPLASMIC.
-!- SIMILARITY: BELONGS TO THE GST SUPERFAMILY. MU FAMILY.
   73; Indels
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             10.8%; Score 468.5; DB 1; Length 43.4%; Pred. No. 9.2e-19; live 36; Mismatches 75; Indels
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         34D8A8FBDD8627C0 CRC64;
 34; Mismatches
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        BY SIMILARITY
                                                                                                                                                                                                                                                                                                                                                                                         PRT;
                                                                                                                                                                                                                                                        182 KKRIEAIPQIDKYLKSSKYIAWPL 205
                                                                                                                                                                                                                                                                                          184 ISRFEGLEKISSYMKSSRFLPKPL 207
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  Biochem. 193:19-24(1990).
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               PRINTS; PR01267; GSTRNSFRASEM.
Transferase; Multigene family.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       217 AA; 25688 MW;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         EMBL; X57489; CAA40726.1; -.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   GST_mu.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        Conservative
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                Query Match
Best Local Similarity 43.48
Walches 89; Conservative
                                                                                                                                                                                                                                                                                                                                                                                       STANDARD;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            InterPro; IPR000521; GST.
InterPro; IPR003081; GST.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          SUBUNIT: HOMODIMER
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    GST; 1.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         S13202; S13202; P04905; 6GST.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     SEQUENCE FROM N.A.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    NCB1_TaxID=10030;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  PF00043;
92;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               fibroblasts.
                                                                                                                                                                                                                                                                                                                                                                                     GTMU_CRILO
000285;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       SEQUENCE
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        INIT_MET
Matches
                                                                                                                                                                                                                                                                                                                                                  RESULT 14
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    Pfam;
                                                                                                                                            qq
                                  ò
                                                                       염
                                                                                                         ô
                                                                                                                                                                                 Qγ
                                                                                                                                                                                                                   qq
                                                                                                                                                                                                                                                        QΥ
                                                                                                                                                                                                                                                                                                                                                                                                    δ
```

Page 11

```
1;
                                                                                                                                             5; Gaps
                                                                                          Query Match 10.8%; Score 468.5; DB 1; Length 217; Best Local Similarity 44.7%; Pred. No. 9.2e-19; Matches 89; Conservative 37; Mismatches 68; Indels 5;
INIT_MET 0 0 BY SIMILARITY. SEQUENCE 217 AA; 25558 MW; C59812A4A30B6375 CRC64;
       FT
```

δy

QQ

Dp ŏ

182 KKRIEAIPQIDKYLKSSKY 200 | | : : | |:|||:: | R4 LARFEGLKKISDYMKSSRF 202 õ

Search completed: January 15, 2002, 14:07:00 Job time: 929 sec

ţ
•

•

,

```
January 15, 2002, 13;49:59; Search time 120.01 Seconds (without alignments) 228.991 Million cell updates/sec
                                                                                                                                                                                                                                                                                                                                                                                                                                                         1 DDHAPREMFDCLDKPIPWRG.....KRLTARKRRSRKEEVEEDEK 371
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      | SIDS8/gcgdata/geneseqp/AA1991.DAT: |
| SIDS8/gcgdata/geneseqp/AA1991.DAT: |
| SIDS8/gcgdata/geneseqp/AA1992.DAT: |
| SIDS8/gcgdata/geneseqp/AA1994.DAT: |
| SIDS8/gcgdata/geneseqp/AA1995.DAT: |
| SIDS8/gcgdata/geneseqp/AA1995.DAT: |
| SIDS8/gcgdata/geneseqp/AA1995.DAT: |
| SIDS8/gcgdata/geneseqp/AA1997.DAT: |
| SIDS8/gcgdata/geneseqp/AA1997.DAT: |
| SIDS8/gcgdata/geneseqp/AA1999.DAT: |
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               /SIDS8/gcgdata/geneseq/geneseqp/AA1988.DAT:*/SIDS8/gcgdata/geneseq/geneseqp/AA1989.DAT:*/SIDS8/gcgdata/geneseq/geneseqp/AA1990.DAT:*
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               /SIDS8/gcgdata/geneseq/geneseqp/AA1981 DAT:*/SIDS8/gcgdata/geneseq/geneseqp/AA1982 DAT:*/SIDS8/gcgdata/geneseq/geneseqp/AA1983 DAT:*/SIDS8/gcgdata/geneseq/geneseqp/AA1983.IDAT:*/SIDS8/gcgdata/geneseq/geneseqp/AA1983.IDAT:*
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         /SIDS8/gcgdata/geneseg/genesegp/AA1985.DAT:*/SIDS8/gcgdata/geneseg/genesegp/AA1986.DAT:*/SIDS8/gcgdata/geneseg/genesegp/AA1987.DAT:*
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        /SIDS8/gcgdata/geneseq/geneseqp/AA1980.DAT:*
GenCore version 4.5
Copyright (c) 1993 - 2000 Compugen Ltd.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             Total number of hits satisfying chosen parameters:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  522463 seqs, 74073290 residues
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries
                                                                                                                                                     OM protein - protein search, using sw model
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    BLOSUM62
Gapop 10.0 , Gapext 0.5
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    Minimum DB seq length: 0
Maximum DB seq length: 200000000
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    A_Geneseq_1101:*
                                                                                                                                                                                                                                                                                                                                                                  US-09-724-296-4
1979
                                                                                                                                                                                                                                                                                                                                                                                                                             Perfect score:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            Scoring table:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                Database :
                                                                                                                                                                                                                                                                                                                                                                                                                                                                     Sequence:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          Searched:
                                                                                                                                                                                                                                        Run on:
                                                                                                                                                                                                                                                                                                                                                                                          Title:
```

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

/SIDS8/gcgdata/geneseg/genesegp/AA2000 .DAT:*/SIDS8/gcgdata/geneseg/genesegp/AA2001 .DAT:*

14: 115: 117: 119: 220:

		dР					
Result No.	Score	Query Match	Query Match Length DB ID	DB	10		Description
1	1979	100.0	!	21	AAY44499 A	AAY44499 A + LOAS CA	S. pombe delta228-
5	1979	100.0		21	AAY44500		GST signal peptide
m	1979	100.0		21	AAY44498 🗝	_	GST signal peptide
4	960	48.5		17	AAW01618		Neurospora crassa
S	910	46.0		21	AAY44502		N. crassa delta228
9	910	46.0		21	AAY44503		B. subtilis delta2
7	230.5			21	AAY44505		D. radiodurans del
80	108.5		388	21	AAG41350		Arabidopsis thalia
6	108.5	5.5	470	21	AAG41349		Arabidopsis thalia
10	108.5	5,5	486	21			Arabidopsis thalia
11	107.5	5.4	615	21			Arabidopsis thalia

opsis the opsis the opsis the opsis the opsis the subunit	π	ppsis ppsis ppsis nce g oolyp oolyp orote n mic	Amino acid sequenc Human Nck associat Shrimp white spot Secreted protein c Human myosin heavy Thermotoga maritim Arabidopsis thalia Human polypeptide Murine Ataxia-tela Human secreted pro
AAG5081 AAG1869 AAG5080 AAG1869 AAG5080	AAR0423 AAG5043 AAG5043 AAG4433 AAG4433 AAG4433	AAG311. AAG311. AAB190. AAB45. AAM400 AAW113 AAB929 AAB190 AAY322	AAW6984 AAY221.5 AAG8493 AAW8560 AAG6456 AAW0160 AAW3013 AAW3617 AAW3617
21 21 21 21 21 21 18	221 221 221 221 221 221 231	221 221 222 221 221 221 221 231	19 20 20 20 20 20 20 20 20 20 20 20 20 20
615 633 662 662 463	8000 463 471 480 321 337 338	20110 20110 20110 329 605 747 747	653 1128 413 677 793 349 1780 3066 215
	7.00000004	444444444	44444444 6.0.0.0.0.0.0.0.0.0.0.0.0.0.0.0.0.0.0.0
107.5 107.5 107.5 107.5 107.5	$\circ \circ $	7.00 7.00	888 800 800 800 800 800 800 800 800 800
12 13 14 15 16	18 20 21 22 23 24	26 27 27 28 30 31 31 32 33 33	33 3 3 3 3 3 3 4 4 4 4 4 4 4 4 4 4 4 4

ALIGNMENTS

```
Delta228-UVDE; ultraviolet damage endonuclease; GST signal peptide; glutathione-S-transferase signal peptide; uvel+ gene product; UV irradiation; DNA damage; UV radiation damage; photoproduct; abasic site; aplatinum diaduct; mismatched nucleotide pairing; nucleotide alkylation; skin cancer.
                                                                                    S. pombe delta228-UV damage endonuclease.
                    AAY44499 standard; Protein; 371 AA
                                                                                                                                                                                                                                                                                                                          Avery AM;
                                                                                                                                                                                                                                                                  98US-0088521.
                                                                                                                                                                                                                                             99WO-US12910.
                                                                                                                                                                                                                                                                            99US-0134752
                                                               (first entry)
                                                                                                                                                                          Schizosaccharomyces pombe
                                                                                                                                                                                                                                                                                                                        Doetsch PW, Kaur B,
                                                                                                                                                                                                                                                                                                                                             WPI; 2000-116417/10.
N-PSDB; AA229859.
                                                                                                                                                                                                                                                                                                   (UYEM-) UNIV EMORY.
                                                                                                                                                                                                 WO9963828-A1.
                                                               27-MAR-2000
                                                                                                                                                                                                                                             08-JUN-1999;
                                                                                                                                                                                                                                                                   08-JUN-1998;
                                                                                                                                                                                                                                                                              18-MAY-1999;
                                                                                                                                                                                                                       16-DEC-1999
                                           AAY44499;
            AAY44499
RESULT
```

```
The present sequence is delta228-UV damage endonuclease. Delta228-UVDE is a 288 amino acid deletion of the N-terminal of the S. pombe uvel+ gene product. This is expressed in frame with a GST leader sequence to generate a fusion protein. This provides stable endonuclease fragments for cleaving a double-stranded DNA molecule that has a distorted structure resulting from UV radiation damage, a photoproduct, an abasic site, mismatched nucleotide pairing, aplatinum diaduct, an intercalated molecule or alkylation of a nucleotide. Uvelp can be used in compositions for internal or topical application and as a therapeutic agent for skin
   A new truncated ultraviolet damage endonuclease for treatment of skin
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                GTQNVLDLIKLVEWNHNFGIHFMRVSSDLFPFASHAKYGYTLEFAQSHLEEVGKLANKYN 120
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                121 HRLTMHPGQYTQIASPREVVVDSAIRDLAYHDEILSRMKLNEQLNKDAVLIIHLGGIFEG 180
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  240
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 301 TMDLMIEAKEKEQAVFELCRRYELQNPPCPLEIMGPEYDQTRDGYYPPGAEKRLTARKRR 360
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    241 TLREGSLDLMPLIPTIRETWTRKGITQKQHYSESADPTAISGMKRRAHSDRVFDFPPCDP 300
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      0; Gaps
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              1 DDHAPREMFDCLDKPIPWRGRLGYACLNTILRSMKERVFCSRTCRITTIQRDGLESVKQL 60
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    GST signal peptide: glutathione-S-transferase signal peptide; ultraviolet damage endonuclease: UVDE: delta228-UVDE; repair-deficient E. coli strain: UV irradiation: DNA damage; UV radiation damage; photoproduct; abasic site; aplatinum diaduct; mismatched nucleotide pairing; nucleotide alkylation;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  181 KKETLDRFRKNYQRLSDSVKARLVLENDDVSWSVQDLLPLCQELNIPLVLDWHHHNIVPG
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   DB 21; Length 371;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  GST signal peptide and delta228 S. pombe UVDE fusion protein.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  Indels
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        100.0%; Score 1979; DB 21; 100.0%; Pred. No. 2.3e-187;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         .;
0
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    100.0%; Pred. no.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       AAY44500 standard; Protein; 600 AA.
                                                                                                  Claim 13; Page 53; 133pp; English.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 fusion protein; skin cancer.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         27-MAR-2000 (first entry)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               Best Local Similarity 100.
Matches 371; Conservative
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             Schizosaccharomyces pombe
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            SRKEEVEEDEK 371
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        361 srkeeveedek 371
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             371 AA;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             Sequence
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             Synthetic
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               Query Match
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               121
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                61
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     361
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       RESULT 2
AAY44500
ID AAY44500
XX AAX4
XC AAX44
XX E GST s
XX GST s
ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            δ
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        ^{0}
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  δ
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                ŏδ
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               q
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     ŏ
```

```
The present sequence is a fusion protein comprising the GST signal peptide and delta228 truncated ultraviolet damage endonuclease (UVDE) from S. pombe. Delta228-UVDE is a 288 amino acid deletion of the N-terminal of the S. pombe uvell gene product. This is expressed in frame with a GST leader sequence. Stable endonuclease fragments can be produced structure resulting from UV radiation damage, a photoproduct, an abasic site, mismatched nucleotide pairing, aplatinum diaduct, an intercalated molecule or alkylation of a nucleotide. Uvelp can be used in compositions for internal or topical application and as a therapeutic agent for skin
                                                                                                                                                                                                                                                                                                                                                A new truncated ultraviolet damage endonuclease for treatment of skin
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    230 ddhapremfdcldkpipwrgrlgyaclntilrsmkervfcsrtcrittigrdglesvkgl 289
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    61 GTQNVLDLIKLVEWNHNFGIHFMRVSSDLFPFASHAKYGYTLEFAQSHLEEVGKLANKYN 120
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            121 HRLTMHPGQYTQIASPREVVVDSAIRDLAYHDEILSRMKLNEQLNKDAVLIIHLGGTFEG 180
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     181 KKETLDRFRKNYQRLSDSVKARLVLENDDVSWSVQDLLPLCQELNIPLVLDWHHHNIVPG 240
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      241 TLREGSLDLMPLIPTIRETWTRKGITQKQHYSESADPTAISGMKRRAHSDRVFDFPPCDP 300
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             Gaps
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           301 TMDLMIEAKEKEQAVFELCRRYELQNPPCPLEIMGPEYDQTRDGYYPPGAEKRLTARKRR 360
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        1 DDHAPREMFDCLDKPIPWRGRLGYACLNTILRSMKERVFCSRTCRITTIQRDGLESVKQL 60
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                100.0%; Score 1979; DB 21; Length 600; 100.0%; Pred. No. 4.8e-187;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 Indels
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               0; Mismatches
               /note= "Encoded by aaaa"
                               /note= "Encoded by gtt" Misc-difference 147
                                                                         /note= "Encoded by gtt"
                                                                                                                                                                                                                                                                                                                                                                                            Disclosure; Page 56; 133pp; English.
                                                                                                                                                                                                                                                                         Avery AM;
                                                                                                                                                                                                             990S-0134752.
                                                                                                                                                                                                 98US-0088521
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             Matches 371; Conservative
                                                                                                                                                                                                                                                                        Kaur B,
                                                                                                                                                                                                                                                                                                  WPI; 2000-116417/10.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     SRKEEVEEDEK 371
                                                                                                                                                                                                                                           (UYEM-) UNIV EMORY.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              Query Match
Best Local Similarity
Misc-difference 11
                                                                                                                                                                                                                                                                                                                N-PSDB; AAZ29860:
                                                                                                       W09963828-A1
                                                                                                                                                                 08-JUN-1999;
                                                                                                                                                                                              08-JUN-1998;
                                                                                                                                                                                                             18-MAY-1999;
                                                                                                                                    16-DEC-1999
                                                                                                                                                                                                                                                                        Doetsch PW,
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       Sednence
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         cancers.
                                                                                                                                                                                                                                                                                                                                                                cancers
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   361
 Οy
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              οy
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         q
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 9
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       Qγ
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 Ωÿ
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             q
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         ŏ
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      qq
```

0

Location/Qualifiers

GST signal peptide; ultraviolet damage endonuclease; UVDE; Uvelp; UV irradiation; DNA damage; UV radiation damage; fusion protein; skin cancer; glutathione-S-transferase. 3ST signal peptide and S. pombe UVDE fusion protein. 1..229 /label= GST_signal_peptide 'note= "Encoded by aaaa" /note= "Encoded by gtt" /note= "Encoded by gtt" note= "S. pombe UVDE" Location/Qualifiers AAY44498 standard; Protein; 828 AA. Claim 7; Page 51; 133pp; English Avery AM 99WO-US12910 98US-0088521 99US-0134752 27-MAR-2000 (first entry) Schizosaccharomyces pombe 230..828 Kaur B, 590 srkeeveedek 600 WPI; 2000-116417/10. N-PSDB; AAZ29858. 82 (UYEM-) UNIV EMORY. Misc-difference Misc-difference Misc-difference WO9963828-A1 08-JUN-1999; 08-JUN-1998; 18-MAY-1999; Doetsch PW, 16-DEC-1999 Synthetic cancers -AAY44498; Peptide Protein AAY44498 g

The present sequence is a fusion protein comprising the GST signal peptide and full length ultraviolet damage endonuclease (UVDE) from Schizosaccharomyces pombe. UVDE genes are isolated by introducing a foreign cDNA library into a repair-deficient E. coli strain and selecting for complemented cells by UV irradiation of the transformants. This provides stable endonuclease fragments for cleaving a double-stranded DNA molecule that has a distorted structure resulting from UV radiation damage, a photoproduct, an abasic site, mismatched nucleotide pairing, aplathium diaduct, an intercalated molecule or alkylation of a nucleotide. Uvelp can be used in compositions for internal or topical application and as a therapeutic agent for skin cancers. new truncated ultraviolet damage endonuclease for treatment of skin 828 AA; Sequence

```
The present sequence shows a DNA repair enzyme which specifically recognises a cyclobutane-type dimer and (6-4)-bound product induced, respectively from TT and TC sequences in DNA, by exposure to UV light. The enzyme cleaves these sequences at the 5'-side. The DNA repair enzyme can be used in, e.g. suntan cream for prevention and treatment of skin cancer. Corventional endonuclease acts on either the cyclobutane-type dimer or the (6-4)-bound product in damaged DNA, while the enzyme of the invention can act on both of them.
                                                                                                                                                         181 KKETLDRFRKNYQRLSDSVKARLVLENDDVSWSVQDLLPLCQELNIPLVLDWHHHNIVPG 240
                                                                                                                                                                                                                TLREGSLDLMPLIPTIRETWTRKGITQKQHYSESADPTAISGMKRRAHSDRVFDFPPCDP 300
                                                                                                                                                                                                                                                                         301 TMDLM1EAKEKEQAVFELCRRYELQNPPCPLETMGPEYDQTRDGYYPPGAEKRLTARKRR 360
                                                                                                                                                                                                                                                                                       DNA repair enzyme; cyclobutane-type dimer; (6-4)-bound product; exposure; UV light; ultraviolet; suntan cream; prevention; treatment;
GTQNVLDLIKLVEWNHNFGIHFMRVSSDLFPFASHAKYGYTLEFAQSHLEEVGKLANKYN 120
                                                        HRLTMHPGQYTQ1ASPREVVVDSAIRDLAYHDEILSRMKLNEQLNKDAVLIIHLGGTFEG 180
                                                                                                              DNA repair enzyme and related DNA - specifically recognises cyclobutane-type dimer and (6-4)-bound product induced by exposure
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  Disclosure; Page 10-13; 18pp; Japanese.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     Neurospora crassa DNA repair enzyme
                                                                                                                                                                                                                                                                                                                                                                                                                                   AAW01618 standard; Protein; 656 AA
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            95JP-0094137
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         95JP-0094137
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            (first entry)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   (SHIS ) SHISEIDO CO LTD.
                                                                                                                                                                                                                                                                                                                                  361 SRKEEVEEDEK 371
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 WPI; 1996-512663/51
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          Neurospora crassa.
                                                                                                                                                                                                                                                                                                                                                              818 srkeeveedek
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                N-PSDB; AAT58286
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     JP08266275-A.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          20-AUG-1997
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            29-MAR-1995;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         29-MAR-1995;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               skin cancer
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                15-0CT-1996
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       to UV light
                                                                                                                                                                                       638
                                           6.1
                                                                                                    121
                                                                                                                                                                                                                                                                                                                                                                                                                      AAW01618
                                                                                                                                                                                                                                                                                                                                                                                                                                   XX
                                                                                                                                                                                                                                                                                                                                                                                                                                                               Dp
                                                                                                                              qq
                                                                                                                                                                                      qq
                                                                                                                                                                                                                                                                         ŏ
                                                                                                                                                                                                                                                                                                      g
                                                                                                                                                                                                                                                                                                                                  ò
                                         Óγ
                                                                       qq
                                                                                                    ò
                                                                                                                                                           δ
```

/ Match 100.0%; Score 1979; DB 21; Length 828; Local Similarity 100.0%; Pred. No. 7.9e-187; nes 371; Conservative 0; Mismatches 0; Indels 0;

Query Match 100. Best Local Similarity 100. Matches 371; Conservative

656 AA;

Sequence

; 0

Gaps

```
10;
                                                                                                                                                                                                106 QSHLEEVGKLANKYNHRLTMHPGQYTQIASPREVVVDSAIRDLAYHDEILSRMKLNEQLN 165
                                                                                                                                                                                                                                        166 KDAVLIIHLGGTFEGKKETLDRFRKNYQRLSDSVKARLVLENDDVSWSVQDLLPLCQELN 225
                                                                                                                                                                                                                                                                                                                                                            282 GMKRRAHSDRVFDFPPCDPTMDLMIEAKEKAREQAVFELCRRYEL-------0 325
                                                                                                                                                                                                                                                                                                                                                                             201 plpwkgrlgyaclntylrnskppifssrtcrmasivdhrhplgfedepehhlknkpdksk 260
                                                                                                                     50 ---QRDGLESVKQLGTQNVLDLIKLVEWNHNFGIHFMRVSSDLFPFASHAKYGYTL-EFA 105
                                                                                                                                                261 epqdelghkfvqelglanardivkmlcwnekygirflrlssemfpfashpvhgyklapfa 320
                                                                                                                                                                                                                                                                                                  226 IPLVLDWHHHNIV--PGTLREGSLDLM--PLIPTIRETWTRKGITQKQHYSESADPTAIS 281
                                                                                                                                                                                                                                                                                                                                                                                                                      326 NPPCP------344
                                                                                                                                                                                                                                                                                                                                                                                                                                                   560 nrpappvkapkkkkggkrkrttdeeaaepeevdtaaddvkdapegpkevpeeeramggpy 619
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       A new truncated ultraviolet damage endonuclease for treatment of skin
                                 Gaps
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      Delta228-UVDE; ultraviolet damage endonuclease; GST signal peptide; glutethione-S-transferase signal peptide; uvel+ gene product; UV irradiation; DNA damage; UV radiation damage; photoproduct; abasic site; aplatinum diaduct; mismatched nucleotide pairing;
                            98;
48.5%; Score 960; DB 17; Length 656; 45.4%; Pred. No. 4.4e-86; Live 54; Mismatches 95; Indels 96
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             crassa delta228-UV damage endonuclease.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 345 ---YYPPGAEKRLTARKRRSRK----EEVEED 369
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              620 nrvywplgceewlkpkkrevkkgkvpeevede 651
                                                         15 PIPWRGRLGYACLNTILRSMKERVFCSRTCRITTI-
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   nucleotide alkylation; skin cancer.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     AAY44502 standard; Protein; 626 AA
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 Avery AM;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        99WO-US12910
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   99US-0134752
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     98US-0088521
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                (first entry)
                              205; Conservative
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              Kaur B,
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         WPI; 2000-116417/10.
              Local Similarity
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                (UYEM-) UNIV EMORY.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   Neurospora crassa.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              WO9963828-A1.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        08-JUN-1999;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            Doetsch PW,
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                27-MAR-2000
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   08-JUN-1998;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     18-MAY-1999;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          16-DEC-1999.
 Query Match
                              Matches
                                                                                       q
                                                                                                                     δ
                                                                                                                                                a
                                                                                                                                                                             Qγ
                                                                                                                                                                                                         Сp
                                                                                                                                                                                                                                         ŏ
                                                                                                                                                                                                                                                                   qq
                                                                                                                                                                                                                                                                                                 ö
                                                                                                                                                                                                                                                                                                                             Op
                                                                                                                                                                                                                                                                                                                                                          ò
                                                                                                                                                                                                                                                                                                                                                                                       qq
                                                                                                                                                                                                                                                                                                                                                                                                                      ŏ
                                                                                                                                                                                                                                                                                                                                                                                                                                                g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                QΥ
```

```
10;
                                                                                                                                                 to generate a fusion protein. This provides stable endonuclease fragments for cleaving a double-stranded DNA molecule that has a distorted structure resulting from UV radiation damage, a photoproduct, an abasic site, mismatched nucleotide pairing, aplatinum diaduct, an intercalated molecule or alkylation of a nucleotide. Uvelp can be used in compositions for internal or topical application and as a therapeutic agent for skin
                                                                               The present sequence is delta228-UV damage endonuclease from N. crassa. Delta228-UVDE is a 288 amino acid deletion of the N-terminal of the uvel+ gene product. This is expressed in frame with a GST leader sequence
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 201 plpwkgrlgyacintylrnakppifssrtcrmasivdhrhplgfedepehhlknkpdksk 260
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            50 ---QRDGLESVKQLGTQNVLDLIKLVEWNHNFGIHFMRVSSDLFPFASHAKYGYTL-EFA 105
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      261 epqdelghkfvqelglanardivkmlcw------fpfashpvhgyklapfa 305
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  106 QSHLEEVGKLANKYNHRLTMHPGQYTQIASPREVVVDSAIRDLAYHDEILSRMKLNEQLN 165
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          306 sevlaeagryaaelghrltthpgqftqlgsprkevvesairdleyhdellsllklpeggn 365
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   226 IPLVLDWHHHNIV--PGTLREGSLDLM--PLIPTIRETWTRKGITQKQHYSESADPTAIS 281
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          Gaps
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       166 KDAVLIIHLGGTFEGKKETLDRFRKNYQRLSDSVKARLVLENDDVSWSVQDLLPLCQELN 225
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      426 ipmvldyhhhnicfdpahlregtldisdpklgeriantwkrkgikqkmhysepcd-gavt 484
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  282 GMKRRAHSDRVFDFPPCDPTMDLMIEAKEKEQAVFELCRRYEL-------Q 325
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         -----PLEIMGPEYDQTRDG----YYPPGAEKRLTA 356
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               Pelta228-UVDE; ultraviolet damage endonuclease; GST signal peptide;
                                                                                                                                                                                                                                                                                                                                                                                                                                           98;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  dlutathione-S-transferase signal peptide; uvel+ gene product;
UV irradiation; DNA damage; UV radiation damage; photoproduct;
abasic site; aplatinum diaduct; mismatched nucleotide pairing;
                                                                                                                                                                                                                                                                                                                                                                                              46.0%; Score 910; DB 21; Length 626;
45.1%; Pred. No. 3.7e-81;
Live 51; Mismatches 91; Indels 99
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                B. subtilis delta228-UV damage endonuclease.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       15 PIPWRGRLGYACLNTILRSMKERVFCSRTCRITTI-
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      nucleotide alkylation; skin cancer.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    AAY44503 standard; Protein; 626 AA.
                                     Claim 16; Page 59; 133pp; English.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              357 RKRRSRK----EEVEED 369
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     605 kkrevkkgkvpeevede 621
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      27-MAR-2000 (first entry)
                                                                                                                                                                                                                                                                                                                                                                                                                     Best Local Similarity 45.13
Matches 197; Conservative
                                                                                                                                                                                                                                                                                                                                 626 AA;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              Bacillus subtilis.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         326 NPPC---
                                                                                                                                                                                                                                                                                                                                   Sequence
                                                                                                                                                                                                                                                                                                                                                                                                   Query Match
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              AAY44503;
                                                                                                                                                                                                                                                                                           cancers.
cancers
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               AAY44503
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            RESULT
                  pp
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  òγ
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       δŏ
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 qq
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              ōλ
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   q
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  QΥ
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 QQ
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                δŏ
```

AAY44505 standard; Protein; 294 AA

```
10;
                                                                                                                                                                                                                                                                           The present sequence is delta228-UV damage endonuclease from B. subtilis. Delta228-UVDE is a 288 amino acid deletion of the N-terminal of the uvel+ gene product. This is expressed in frame with a GST leader sequence to generate a fusion protein. This provides stable endonuclease fragments for cleaving a double-stranded DNA molecule that has a distorted
                                                                                                                                                                                                                                                                                                                                                 structure resulting from UV radiation damage, a photoproduct, an abasic site, mismatched nucleotide pairing, aplatinum diaduct, an intercalated molecule or alkylation of a nucleotide. Uvelp can be used in compositions for internal or topical application and as a therapeutic agent for skin
                                                                                                                                                                                                             skin
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            ---QRDGLESVKQLGTQNVLDLIKLVEWNHNFGIHFMRVSSDLFPFASHAKYGYTL-EFA 105
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    261 epqdelghkfvqelglanardivkmlcw--------fpfashpvhgyklapfa 305
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      KDAVLIIHLGGTFEGKKETLDRFRKNYQRLSDSVKARLVLENDDVSWSVQDLLPLCQELN 225
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            GMKRRAHSDRVFDFPPCDPTMDLMIEAKEKEQAVFELCRRYEL-------Q 325
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  326 NPPC-----YYPPGAEKRLTA 356
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           Gaps
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             226 IPLVLDWHHHNIV--PGTLREGSLDLM--PLIPTIRETWTRKGITQKQHYSESADPTAIS
                                                                                                                                                                                                           new truncated ultraviolet damage endonuclease for treatment of
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          98;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                DB 21; Length 626;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          Indels
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          91;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                             46.0%; Score 910; DB 21; 45.1%; Pred. No. 3.7e-81;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       45.1%; Pred. ...
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   15 PIPWRGRLGYACLNTILRSMKERVFCSRTCRITI
                                                                                                                                                                                                                                                   Claim 16; Page 59; 133pp; English.
                                                                                                                                                     Avery AM;
                                                    99WO-US12910.
                                                                                98US-0088521
99US-0134752
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  :11 :1 iiii...
kkrevkkgkvpeevede 621
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  357 RKRRSRK----EEVEED 369
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          Best Local Similarity 45.1%
Matches 1.97; Conservative
                                                                                                                                                     Kaur B,
                                                                                                                                                                                WPI; 2000-116417/10
                                                                                                                          (UYEM-) UNIV EMORY.
                                                                                                                                                                                                                                                                                                                                                                                                                                    626 AA;
                                                                                08-JUN-1998;
W09963828-A1
                           16-DEC-1999
                                                                                                                                                     Doetsch PW,
                                                                                                                                                                                                                                                                                                                                                                                                                                      Sequence
                                                                                                                                                                                                                                                                                                                                                                                                                                                                             Query Match
                                                                                                                                                                                                                          cancers
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   282
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  366
 qq
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         q
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   à
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               à
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           ŏ
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          õ
```

```
The present sequence is delta228-UV damage endonuclease from D. radiodurans. Delta228-UVDE is a 288 amino acid deletion of the N-terminal of the uvel+ gene product. This is expressed in frame with a GST leader sequence to generate a fusion protein. This provides stable endonuclease fragments for cleaving a double-stranded DNA molecule that has a distorted structure resulting from UV radiation damage, a photoproduct, an abasic site, mismatched nucleotide pairing, aplatinum diaduct, an intercalated molecule or alkylation of a nucleotide. Uvelp can be used in compositions for internal or topical application and as a
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                A new truncated ultraviolet damage endonuclease for treatment of skin
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       218 gedpsvrewvlraratwgppewgvvhls----ngiegpgdrrhshliadfpsayadvpg 272
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         59 rlyrlssslfpmldlagddtgaavlthlapqlleaghaftdagvrllmhpeqfivlnsdr 118
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           81 HFMRVSSDLFP---FASHAKYGYTLEFAQSHLEEVGKLANKYNHRLTMHPGQYTQIASPR 137
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          193 QRLSDSVKARLVLENDDVSWSVQDLLPLCQELNIPLVLDWHHHNIVPGTLREGSLDLMP- 251
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 252 -LIPTIRETWTRKGIT-----QKQHYSESADPTAISGMKRRAHSDRVFDFPPCDPTM-D 303
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              21 RLGYACL.NTILLRSMKERVFCSRTCRITTIQRDGLESVKQLGTONVLDLIKLVEWNHNFGI 80
                                                                                                                                           Delta228-UVDE; ultraviolet damage endonuclease; GST signal peptide; dutathione-S-transferase signal peptide; uval+ gene product; UV irradiation; DNA damage; UV radiation damage; photoproduct; abasic site; aplatinum diaduct; mismatched nucleotide pairing;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 1 qlglvcltvgpevrfrtvtlsryralspaerea--klldlyssniktlrgaadycaahdi
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                138 EVVVDSAIRDLAYHDEI-----LSRMKLNEQLNKDAVLIIHLGGTFEGKKETLDRFRKNY
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       DB 21; Length 294;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   11.6%; Score 230.5; DB 21; Length; 27.9%; Pred. No. 2.3e-14; ive 42; Mismatches 140; Indels
                                                                                                      D. radiodurans delta228-UV damage endonuclease.
                                                                                                                                                                                                                       nucleotide alkylation; skin cancer.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            therapeutic agent for skin cancers
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    Claim 16; Page 60; 133pp; English.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      Avery AM;
                                                                                                                                                                                                                                                                                                                                                                                                         98US-0088521
99US-0134752
                                                                                                                                                                                                                                                                                                                                                                      99WO-US12910
                                                                     (first entry)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      Best Local Similarity 27.9%
Matches 88; Conservative
                                                                                                                                                                                                                                                          Dienococcus radiodurans.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      Doetsch PW, Kaur B,
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           WPI; 2000-116417/10
                                                                                                                                                                                                                                                                                                                                                                                                                                                                 (UYEM-) UNIV EMORY.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                294 AA;
                                                                                                                                                                                                                                                                                                                                                                      08-JUN-1999;
                                                                                                                                                                                                                                                                                                                                                                                                                          18-MAY-1999;
                                                                     27-MAR-2000
                                                                                                                                                                                                                                                                                            W09963828-A1
                                                                                                                                                                                                                                                                                                                                 16-DEC-1999
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       Query Match
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                Sequence
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   cancers
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            167
δy
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   d
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         qq
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   qq
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          q
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        δλ
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                δ
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   ó
```

10;

Gaps

AAY44505 RESULT

```
| 17 - JUN - 1999;
| 18 - JUN - 1999;
                                                                                                                                                                                                                                                                                                                                                                 08-JUL-1999;
09-JUL-1999;
12-JUL-1999;
13-JUL-1999;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     05-AUG-1999;
06-AUG-1999;
06-AUG-1999;
09-AUG-1999;
09-AUG-1999;
                                                                                                                                                                                                                                                                                            30-JUN-1999;
01-JUL-1999;
01-JUL-1999;
02-JUL-1999;
06-JUL-1999;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             04-AUG-1999;
04-AUG-1999;
05-AUG-1999;
                                                                                                                                                                                                                                                  1999;
1999;
1999;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 999;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       16661
16861
                                                                                                                                                                                                                                                                                                                                                                                                                             :666
                                                                                                                                                                                                                                                                                                                                                                                                                                                         :6661
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      16661
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  16661
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            :6661
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             4666
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         :6661
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       :6661
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               23-JUL-1999;
26-JUL-1999;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              16661
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             1999;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  16661
                                                                                                                                                                                                                       1999;
                                                                                                                                                                                                                                      1999;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                        16661
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             16661
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      28-JUL-1999;
02-AUG-1999;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            12-AUG-1999;
13-AUG-1999;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   02-AUG-19
02-AUG-19
03-AUG-19
                                                                                                                                                                                                                                                                                                                                                                                                                                                       6-JUL-1
6-JUL-1
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     19-JUL-1
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 21-JUL-19
21-JUL-19
21-JUL-19
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         27 - JUL-19
27 - JUL-19
                                                                                                                                                                                                                                                                                                                                                                                                                             4-JUL-1
5-JUL-1
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              9-JUL-1
9-JUL-1
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         2-JUL-1
                                                                                                                                                                                                                                                 24 - JUN-
28 - JUN-
29 - JUN-
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   9-JUL-1
                                                                                                                                                                                                                       23-JUN-
      Protein identification; signal transduction pathway; metabolic pathway; hybridisation assay; genetic mapping; gene expression control; promoter; termination sequence.
                                                                                                                                                                                                       Arabidopsis thaliana protein fragment SEQ ID NO: 51435.
                                                                                                                   AAG41350 standard; Protein; 388 AA
                                                                                                                                                                                                                                                                                                                                                                                                           990S-0121825
990S-0123180
990S-0125788
990S-0125788
990S-0125788
990S-0127462
990S-0127462
990S-0127462
990S-0127462
990S-0127462
990S-0127462
990S-013748
990S-013248
990S-0132486
990S-0132486
990S-0132486
990S-0132486
990S-0134219
990S-0134219
990S-0134221
990S-0136021
990S-0136021
990S-0136021
990S-0136021
990S-0137502
990S-0137502
                                                                                                                                                                                                                                                                                                                                                                                 25-FEB-2000; 2000EP-0301439
                                                                                                                                                                            18-OCT-2000 (first entry)
               304 LMIEAKEKEQAVFEL 318 : 1|| ||:|: | 273 ieveakgkeeaiaal 287
                                                                                                                                                                                                                                                                                              Arabidopsis thaliana.
                                                                                                                                                                                                                                                                                                                        EP1033405-A2.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              14 - MAY - 1999;
14 - MAX - 1999;
14 - MAX - 1999;
18 - MAX - 1999;
19 - MAX - 1999;
20 - MAX - 1999;
21 - MAX - 1999;
22 - MAX - 1999;
25 - MAX - 1999;
26 - MAX - 1999;
27 - MAX - 1999;
28 - MAX - 1999;
29 - MAX - 1999;
30 - JUN - 1999;
30 - JUN - 1999;
31 - JUN - 1999;
                                                                                                                                                                                                                                                                                                                                                                                                                       05-MAR-1999
09-MAR-1999
23-MAR-1999
25-MAR-1999
06-APR-1999
06-APR-1999
19-APR-1999
19-APR-1999
23-APR-1999
23-APR-1999
23-APR-1999
23-APR-1999
06-MAY-1999
06-MAY-1999
06-MAY-1999
06-MAY-1999
06-MAY-1999
                                                                                                                                                                                                                                                                                                                                                     06-SEP-2000.
                                                                                                                                                                                                                                                                                                                                                                                                              25-FEB-1999,
                                                                                                                                               AAG41350;
                                                                                         œ
                                                                                                       AAG41350
                                                                                         RESULT
                                                                                                                      THE STANTANT OF THE STANTANT O
                  ò
                                            PP
```

```
9905 - 0139454
9905 - 0139455
9905 - 0139457
9905 - 0139458
9905 - 0139459
9905 - 0139460
9905 - 0139461
9905 - 0139461
9905 - 0139463
9905 - 0139750
9905 - 0139817
9905 - 0139817
                                                                                                                                                                                                                                                                          9905-0140695.
9905-0140823.
9905-0140823.
9905-0141287.
9905-0142154.
9905-0142055.
9905-0142305.
9905-0142305.
9905-0142303.
                                                                                                                                                                                                                                                                                                                                                                                                                                                      990S-0143542.
990S-0143624.
990S-0144005.
990S-0144085.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       990S - 0144325 - 990S - 0144331 - 990S - 0144331 - 990S - 0144333 - 990S - 0144334 - 990S - 0144335 - 990S - 0144335 - 990S - 0144335 - 990S - 0144335 - 990S - 014508 - 990S - 014708 - 990S - 014708 - 990S - 0147192 - 990S - 0147192 - 990S - 0147193 - 990S - 0148171 - 999S - 0148319 - 990S - 0148319
99US-0139492
```

```
Protein identification; signal transduction pathway; metabolic pathway; hybridisation assay; genetic mapping; gene expression control; promoter; termination sequence.
                                                            211 tiyeragriegrkgsitgi-----piltmpnddithptpdltgyiteggiyld 258
                                                                                                                 170 LIIHLGGTFEGGRETLDRFRKNYQRLSDSVKARLVLENDDVSWSVQDLLPLCQELNIPLV 229
                                                                                                                                                                                       230 LDWHHHNIVPGTLREGSLDLMPLIPTIRETWTRKGITQKQH-----YSESADPTAISG 282
                                                                                                                                                                                                                  259 rqlhnrqiyp-----pinvlpslsrlmksaigegmtrkdhsdvsnqlyanyaigkdvqa 312
     101 gednfalvťaamgvnm-----etaqffkrdfeengsmervtlflnlandptieri 150
                                       119 YNHRLTMHPGQYTQIASPREVVV-----DSAIRDL-AYHDEILSRMKLNEQLNKDAV 169
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     Arabidopsis thallana protein fragment SEQ ID NO: 51434.
                                                                                                                                                                                                                                                                                  283 MKRRAHSDRVFDFPPCDPTMDLM-1EAKEKEQAVFELCRRYELQN 326
                                                                                                                                                                                                                                                                                                                                                                                           AAG41349 standard; Protein; 470 AA.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            99US-0132487.
99US-0132863.
99US-0134256.
99US-0134218.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              99US-0123180.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     99US-0126785.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               99US-0129845.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 99US-0130891.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     99US-0132048.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         99US-0132484.
99US-0132485.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             99US-0132486.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         25-FEB-2000; 2000EP-0301439.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   99US-0126264
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            99US-0128714
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               99US-0130449
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  99US-0130510.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      99US-0134219
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          99US-0134370.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         99US-0134768.
99US-0134941.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  99US-0125788
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          99US-0128234
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              99US-0130077
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       99US-0134221
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             99US-0135124
99US-0135353
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 99US-0135629
                                                                                                                                                                                                                                                                                                                                                                                                                                                                 18-OCT-2000 (first entry)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 Arabidopsis thaliana,
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    01-APR-1999;
06-APR-1999;
08-APR-1999;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   EP1033405-A2.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 23-MAR-1999;
25-MAR-1999;
29-MAR-1999;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             06-MAY-1999;
06-MAY-1999;
07-MAY-1999;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      30-APR-1999;
04-MAY-1999;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                11-MAY-1999;
14-MAY-1999;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       06-SEP-2000,
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               .6-APR-1999;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              19-APR-1999;
21-APR-1999;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  .APR-1999;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 23-APR-1999;
28-APR-1999;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       30-APR-1999;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          05-MAY-1999;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      14-MAY-1999;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     14-MAY-1999;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          .4-MAY-1999;
                                                                                                                                                                                                                                                                                                                                                                                                                               AAG41349;
                                                                                                                                                                                                                                                                                                                                                                        AAG41349
                                                                                                                                                                                                                                                                                                                                                                                                                                 qq
                                                                                                                                                   qq
                                                                                                                                                                                                                          qq
                                       Ωý
                                                                                                                 Öλ
                                                                                                                                                                                       ò
                                                                                                                                                                                                                                                               ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    16;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          Ouery Match 5.5%; Score 108.5; DB 21; Length 388; Best Local Similarity 20.0%; Pred. No. 0.042; Matches 69; Conservative 60; Mismatches 123; Indels 93; Gaps
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         ---NFGIHFMRVSSDLFPFASHAKYGYTLEFAQSHLEEVG-----KLANK---- 118
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    LNTILRSMKERVF------CSRTCRITTIQRDGLESVKQLGTQNVLDLIKLVEWNH 76
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   99US-0159330.
99US-0159331.
99US-0159637.
99US-0159584.
99US-0159584.
990S-0148684.
990S-0149178.
990S-0149175.
990S-0149722.
990S-0149923.
990S-0149923.
990S-0149930.
990S-015984.
990S-0150864.
990S-0151065.
990S-0151303.
990S-0151303.
990S-0151303.
                                                                                                                                                                                                                                                                                                                                                                  99US - 0154018
99US - 0154039
99US - 0154779
99US - 0155139
99US - 0155659
99US - 0156596
99US - 0156753
99US - 0157753
99US - 0157753
99US - 0157753
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            99US-0159293.
99US-0159294.
99US-0159295.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                99US-0160768.
99US-0160770.
99US-0160814.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      99US-0161992.
99US-0161993.
99US-0162142.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     99US-0160815.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         99US-0160981.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            99US-0161405.
99US-0161406.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  90S-0161359
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 99US-0161360
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     99US-0161920
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               99US-0160767
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               99US-0161404
                                                   18-AG-1999;
20-AUG-1999;
20-AUG-1999;
20-AUG-1999;
23-AUG-1999;
                                                                                                                                                                                                                  27 - AUG - 1999;
27 - AUG - 1999;
30 - AUG - 1999;
31 - AUG - 1999;
01 - SEP - 1999;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           5-OCT-1999;
5-OCT-1999;
5-OCT-1999;
6-OCT-1999;
6-OCT-1999;
                                                                                                                                                                                  -AUG-1999;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                :666
                                                                                                                                              23-AUG-1999;
25-AUG-1999;
                                                                                                                                                                                                                                                                                                                                                                                                        :666
                                                                                                                                                                                                                                                                                                                                  -SEP-1999;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     999
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    06-0CT-19
07-0CT-19
08-0CT-19
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               13 - OCT - 19
14 - OCT - 19
18 - OCT - 19
21                                                                                                                                                                                                                                                                                                                                                                                                                                                              -SEP-1
-SEP-1
                                                                                                                                                                                                                                                                                                                                                                                       -SEP-1
                                                                                                                                                                                                                                                                                                                                                                                                                           -SEP-1
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  -0CT-1
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       -0CT-1
                                                                                                                                                                                                                                                                                                                                                                        SEP-1
                                                                                                                                                                                                                                                                                                                                                                                                          ·SEP-1
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     -0CT-1
                                                                                                                                                                                                                                                                                                                                                                                                                                              SEP-1
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      27
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            11
```

oy Oy us-09-724-296-4.rag

Page 8

99US-0147 99US-0147 99US-0147	99US-0147	990S-0147	990S-0147	99US-0148 99US-0148	99US-0148	990S-0148	99US-0149 99US-0149	99US-0149	990S-0149	99US-0149	99US-0149	990S-0150	99US-0151	99US-0151	990S-0151	990S-0151	99US-0152	9908-0153	99US-0154	99US-0154	99US-0155	9908-0155	990s-0156	99US-0156 99US-0157	99US-0157	990S-0158	99US-0158	990s-0159	99US-0159	99US-0159	99US-0159	99US-0159	99US-0159	99US-0160	990S-0160	990S-0160	990S-0160	99US-0160	99US-0160	990S-016.	990S-016	99US-0161 99US-0161	99US-0161	990S-016
03-AUG-1999; 04-AUG-1999; 04-AUG-1999;	US-AUG-1999; 05-AUG-1999; 06-AUG-1999;	06-AUG-1999;	09-AUG-1999; 09-AUG-1999;	10-AUG-1999; 11-AUG-1999;	1.2-AUG-1999;	13-AUG-1999; 13-AUG-1999;	16-AUG-1999; 17-AUG-1999;	18-AUG-1999;	20-AUG-1999; 20-AUG-1999;	20-AUG-1999; 23-AUG-1999;	23-AUG-1999;	25-AUG-1999; 26-AUG-1999;	27-AUG-1999;	27-AUG-1999; 27-AUG-1999;	30-AUG-1999;	31-AUG-1999; 01-SEP-1999;	07-SEP-1999;	10-SEP-1999; 13-SEP-1999;	15-SEP-1999;	16-SEP-1999;	22-SEP-1999;	23-SEP-1999;	28-SEP-1999;	29-SEP-1999; 04-OCT-1999;	05-OCT-1999;	07-0CT-1999;	08-OCT-1999;	13-0CT-1999;	13-OCT-1999; 13-OCT-1999;	14-OCT-1999;	14-0CT-1999; 14-0CT-1999;	14-OCT-1999;	18-OCT-1999;	21-OCT-1999; 21-OCT-1999;	21-OCT-1999;	21-OCT-1999; 21-OCT-1999:	21-OCT-1999;	22-OCT-1999; 22-OCT-1999;	22-0CT-1999;	25-0CT-1999; 25-0CT-1999;	25-0CT-1999;	26-OCT-1999; 26-OCT-1999;	26-OCT-1999;	28-0CT-1999; 28-0CT-1999;
7 7 7 7 8 8 8 8 8 8 8 8 8 8 8 8 8 8 8 8																																												
																	·	·																										
99US-0136021. 99US-0136392. 99US-0136782.	99US-0137222. 99US-0137528.	990S-0137724.	99US-0138094. 99US-0138540.	99US-0138847. 99US-0139119.	99US-0139452.	990S-0139493.	99US-0139454. 99US-0139455.	990S-0139456.	99US-0139457. 99US-0139458.	99US-0139459. 99US-0139460.	99US-0139461.	99US-0139462. 99US-0139463.	99US-0139750.	99US-0139763. 99US-0139817.	990S-0139899.	990S-0140353. 990S-0140354.	99US-0140695.	990S-0140823.	99US-0141287.	99US-0141842.	99US-0142055.	99US-0142390.	990S-0142920.	99US-0142977. 99US-0143542.	99US-0143624.	99US-0144085.	99US-0144086.	990S-0144331.	99US-0144332. 99US-0144333	99US-0144334.	99US-0144335. 99US-0144352.	99US-0144632.	99US-0144814.	99US-0145086. 99US-0145088	99US-0145085.	99US-0145087.	99US-0145192.	99US-0145145.	99US-0145224.	99US-0145276.	99US-0145918.	99US-0145919. 99US-0145951.	99US-0146386.	990S-0146389.
25-MAY-1999; 27-MAY-1999; 28-MAY-1999;	03 - JUN - 1999; 03 - JUN - 1999;	07-JUN-1999;	08-JUN-1999; 10-JUN-1999;	10-JUN-1999; 14-JUN-1999;	16-JUN-1999;	17-JUN-1999;	18-JUN-1999; 18-JUN-1999;	18-JUN-1999;	18-JUN-1999;	18-JUN-1999; 18-JUN-1999;	18-JUN-1999;	18-JUN-1999; 18-JUN-1999;	18-JUN-1999;	18-JUN-1999; 21-JUN-1999;	22-JUN-1999;	23-JUN-1999;	24 - JUN-1999;	28-JUN-1999;	30-JUN-1999;	01-JUL-1999;	02-JUL-1999;	06-JUL-1999;	09-JUL-1999;	12-JUL-1999; 13-JUL-1999;	14-JUL-1999;	16-JUL-1999;	16-JUL-1999;	19-JUL-1999;	19-JUL-1999; 19-JUL-1999;	19-JUL-1999;	19-JUL-1999; 20-JUL-1999;	20-JUL-1999;	21-JUL-1999;	21-JUL-1999; 21-JUL-1999;	22-JUL-1999;	22-JUL-1999; 22-JUL-1999;	22-JUL-1999;	23-JUL-1999; 23-JUL-1999;	23-JUL-1999;	26-JUL-1999;	27 - JUL - 1999;	27-JUL-1999; 28-JUL-1999;	02-AUG-1999;	02-AUG-1999;
7 7 7 8 8 8 8 8 8 8 8 8 8 8 8 8 8 8 8 8																																												•

447038 472034 4772034 4772034 4772036 477206 477206 477206 477

```
18-MAY-1999;
19-MAY-1999;
20-MAY-1999;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            1999;
                                                                                                                                                                                                                               27-MAY-1999;
28-MAY-1999;
                                                                                                                                                                                                                                                                                                                                             10-JUN-1999;
10-JUN-1999;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  1999;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               :6661
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         666;
                                                                                                                                                                                               24-MAY-1999;
25-MAY-1999;
                                                                                                                                                                                                                                                                              03-JUN-1999;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 1999;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   29-JUN-1999;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          6-JUL-1999;
                                                                                                                                                                                  MAY-1999;
                                                                                                                                                                                                                                                                 -JUN-1999;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       20-JUL-
        16;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           Protein identification; signal transduction pathway; metabolic pathway; hybridisation assay; genetic mapping; gene expression control; promoter; termination sequence.
                                                                                                                                :|| | | :| : || : : || : :|| : :|| : :|| : :|| :|| :|| :|| :|| :|| :|| :|| :|| :|| :|| :|| :|| :|| :|| :|| :|| :|| :|| :|| :|| :|| :|| :|| :|| :|| :|| :|| :|| :|| :|| :|| :|| :|| :|| :|| :|| :|| :|| :|| :|| :|| :|| :|| :|| :|| :|| :|| :|| :|| :|| :|| :|| :|| :|| :|| :|| :|| :|| :|| :|| :|| :|| :|| :|| :|| :|| :|| :|| :|| :|| :|| :|| :|| :|| :|| :|| :|| :|| :|| :|| :|| :|| :|| :|| :|| :|| :|| :|| :|| :|| :|| :|| :|| :|| :|| :|| :|| :|| :|| :|| :|| :|| :|| :|| :|| :|| :|| :|| :|| :|| :|| :|| :|| :|| :|| :|| :|| :|| :|| :|| :|| :|| :|| :|| :|| :|| :|| :|| :|| :|| :|| :|| :|| :|| :|| :|| :|| :|| :|| :|| :|| :|| :|| :|| :|| :|| :|| :|| :|| :|| :|| :|| :|| :|| :|| :|| :|| :|| :|| :|| :|| :|| :|| :|| :|| :|| :|| :|| :|| :|| :|| :|| :|| :|| :|| :|| :|| :|| :|| :|| :|| :|| :|| :|| :|| :|| :|| :|| :|| :|| :|| :|| :|| :|| :|| :|| :|| :|| :|| :|| :|| :|| :|| :|| :|| :|| :|| :|| :|| :|| :|| :|| :|| :|| :|| :|| :|| :|| :|| :|| :|| :|| :|| :|| :|| :|| :|| :|| :|| :|| :|| :|| :|| :|| :|| :|| :|| :|| :|| :|| :|| :|| :|| :|| :|| :|| :|| :|| :|| :|| :|| :|| :|| :|| :|| :|| :|| :|| :|| :|| :|| :|| :|| :|| :|| :|| :|| :|| :|| :|| :|| :|| :|| :|| :|| :|| :|| :|| :|| :|| :|| :|| :|| :|| :|| :|| :|| :|| :|| :|| :|| :|| :|| :|| :|| :|| :|| :|| :|| :|| :|| :|| :|| :|| :|| :|| :|| :|| :|| :|| :|| :|| :|| :|| :|| :|| :|| :|| :|| :|| :|| :|| :|| :|| :|| :|| :|| :|| :|| :|| :|| :|| :|| :|| :|| :|| :|| :|| :|| :|| :|| :|| :|| :|| :|| :|| :|| :|| :|| :|| :|| :|| :|| :|| :|| :|| :|| :|| :|| :|| :|| :|| :|| :|| :|| :|| :|| :|| :|| :|| :|| :|| :|| :|| :|| :|| :|| :|| :|| :|| :|| :|| :|| :|| :|| :|| :|| :|| :|| :|| :|| :|| :|| :|| :|| :|| :|| :|| :|| :|| :|| :|| :|| :|| :|| :|| :|| :|| :|| :|| :|| :|| :|| :|| :|| :|| :|| :|| :|| :|| :|| :|| :|| :|| :|| :|| :|| :|| :|| :|| :|| :|| :|| :|| :|| :|| :|| :|| :|| :|| :|| :|| :|| :|| :|| :|| :|| :|| :|| :|| :|| :|| :|| :|| :|| :|| :|| :|| :|| :|| :|| :|| :|| :|| :|| :|| :|| :|| :|| :|| :|| :|| :|| :|| :|| :|| :|| :|| :|| :|| :|| :|| :|| :|| :|| :|| :|| :|| :|| :|| :|| :|| :|| :|| :|| :|| :|| :|| :|| :||
                                                                                                                                                                          77 ---NFGIHFMRVSSDLFPFASHAKYGYTLEFAQSHLEEVG------KLANK---- 118
                                                                                                                                                                                               119 YNHRLTMHPGQYTQIASPREVVV------DSAIRDL-AYHDEILSRMKLNEQLNKDAV 169
                                                                                                                                                                                                                                                             170 LIIHLGGTFEGKKETLDRFRKNYQRLSDSVKARLVLENDDVSWSVQDLLPLCQELNIPLV 229
                                                                                                                                                                                                                                                                                                                           230 LDWHHHNIVPGTLREGSLDLMPLIPTIRETWTRKGITQKQH-----YSESADPTAISG 282
                                                                                                                                                                                                                                                                                                                                                                                            Gaps
                                                                                                            27 LNTILRSMKERVF-----CSRTCRITTIQRDGLESVKQLGTQNVLDLIKLVEWNH 76
                                                                                   93;
                                             / Match 5.5%; Score 108.5; DB 21; Length 470; Local Similarity 20.0%; Pred. No. 0.056; les 69; Conservative 60; Mismatches 123; Indels 93;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             Arabidopsis thaliana protein fragment SEQ ID NO: 51433.
                                                                                                                                                                                                                                                                                                                                                                                                                                    283 MKRRAHSDRVFDFPPCDPTMDLM-IEAKEKEQAVFELCRRYELQN 326
                                                                                                                                                                                                                                                                                                                                                                                                                                                       AAG41348 standard; Protein; 486 AA
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    990S-0121825.
990S-0123180.
990S-0125748.
990S-0126264.
990S-0126785.
990S-0126785.
990S-0128714.
990S-0128714.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                990S-0130891.
990S-0131449.
990S-0132048.
990S-0132407.
990S-0132484.
 99US-0161993.
99US-0162142.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      25-FEB-2000; 2000EP-0301439
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    99US-0130510
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             18-OCT-2000 (first entry)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            Arabidopsis thaliana.
28-OCT-1999;
29-OCT-1999;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        EP1033405-A2.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    25 - FEB - 1999;
05 - MAR - 1999;
23 - MAR - 1999;
25 - MAR - 1999;
26 - MAR - 1999;
01 - APR - 1999;
06 - APR - 1999;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                16-APR-1999;
19-APR-1999;
21-APR-1999;
23-APR-1999;
23-APR-1999;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              30-APR-1999;
04-MAY-1999;
05-MAY-1999;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        06-SEP-2000
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                AAG41348;
                                                  Query Match
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   RESULT 10
                                                                              Matches
 PR
PR
                                                                                                                                                                             ŏ
                                                                                                                                                                                                          q
                                                                                                                                                                                                                                         ò
                                                                                                                                                                                                                                                                    qq
                                                                                                                                                                                                                                                                                                       ò
                                                                                                                                                                                                                                                                                                                                  Q
                                                                                                                                                                                                                                                                                                                                                                     δ
                                                                                                                                                                                                                                                                                                                                                                                                   9
                                                                                                                                                                                                                                                                                                                                                                                                                                    à
                                                                                                                                                                                                                                                                                                                                                                                                                                                                 g
```

```
990S-0132863.
990S-0132186.
990S-0134218.
990S-0134221.
990S-0134221.
990S-0134370.
990S-013434941.
                                                                                                                     99US-0135353.
99US-0135629.
99US-0136021.
                                                                                                                                                   990S-0136392-
990S-0136782-
990S-013722-
990S-0137528-
990S-0137724-
990S-0137724-
                                                                                                                                                                                                                             9905 - 0138540
9905 - 0138847
9905 - 0139453
9905 - 0139453
9905 - 0139454
9905 - 0139456
9905 - 0139456
9905 - 0139456
9905 - 0139460
9905 - 0139461
9905 - 0139461
                                                                                                                                                                                                                                                                                                                                                                                                                    990S-0139763.
990S-0139817.
990S-0139899.
                                                                                                                                                                                                                                                                                                                                                                                                                                                               990S-0140354
990S-0140695
990S-0140891
990S-0140991
990S-0142187
990S-0142184
990S-0142184
990S-0142184
990S-0142803
990S-0142803
990S-014384
990S-0143624
990S-0143624
990S-0143624
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     99US-0144325.
99US-0144331.
99US-0144332.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            99US-0145088.
99US-0145085.
99US-0145087.
99US-0132486.
99US-0132487.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 99US-0144334.
99US-0144335.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        99US-0144352
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 99US-0144632
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       99US-0144814
```

us-09-724-296-4.rag

```
99US-0150566.
99US-0150884.
99US-0151065.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   990S-0153758.
990S-0154018.
990S-0154039.
990S-0154779.
990S-0155139.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    990S-0155659.
990S-0156458.
990S-0156596.
990S-0157117.
990S - 0145089 .
990S - 0145185 .
990S - 0145145 .
990S - 0145218 .
990S - 0145214 .
990S - 0145913 .
990S - 0145918 .
990S - 0145918 .
990S - 0145918 .
990S - 0146386 .
990S - 0146386 .
990S - 0146386 .
990S - 0146386 .
990S - 0147303 .
990S - 0147303 .
                                                                                                                                                                                                                                                                                                                                                                  990S - 0.148171
990S - 0.148319
990S - 0.148365
990S - 0.148565
990S - 0.149368
990S - 0.149368
990S - 0.149175
990S - 0.149722
990S - 0.149929
990S - 0.149929
990S - 0.149929
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    990S-0151066.
99US-0151080.
99US-0151303.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    99US-0151438.
99US-0151930.
99US-0152363.
99US-0153070.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      99US-0159638.
99US-0159584.
99US-0160741.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      99US-0160767.
99US-0160768.
99US-0160770.
                                                                                                                                                                                                                                                                                                                   99US-0147416.
99US-0147493.
99US-0147935.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       99US-0159330.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        99US-0157865.
99US-0158029.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         99US-0158232.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       99US-0158369.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          99US-0159294.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        99US-0159295
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          99US-0159637
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         59329
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         99US-01
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      21-OCT-1999;
21-OCT-1999;
21-OCT-1999;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  20 - AuG - 1999
20 - AuG - 1999
23 - AuG - 1999
23 - AuG - 1999
25 - AuG - 1999
27 - AuG - 1999
27 - AuG - 1999
27 - AuG - 1999
31 - AuG - 1999
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     16-SEP-1999;
22-SEP-1999;
23-SEP-1999;
24-SEP-1999;
26-SEP-1999;
05-OCT-1999;
06-OCT-1999;
07-OCT-1999;
08-OCT-1999;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     12-071-1999
13-077-1999
13-077-1999
14-077-1999
14-077-1999
14-077-1999
14-077-1999
14-077-1999
18-077-1999
21-077-1999
22 - JUL - 1999;
23 - JUL - 1999;
23 - JUL - 1999;
23 - JUL - 1999;
24 - JUL - 1999;
27 - JUL - 1999;
28 - JUL - 1999;
28 - JUL - 1999;
29 - AUG - 1999;
40 - AUG - 1999;
60 - AUG - 1999;
61 - AUG - 1999;
62 - AUG - 1999;
63 - AUG - 1999;
64 - AUG - 1999;
65 - AUG - 1999;
66 - AUG - 1999;
66 - AUG - 1999;
67 - AUG - 1999;
68 - AUG - 1999;
69 - AUG - 1999;
61 - A
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      07-SEP-1999;
10-SEP-1999;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       15-SEP-1999;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        18-AUG-1999
```

```
16;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   Protein identification; signal transduction pathway; metabolic pathway; hybridisation assay; genetic mapping; gene expression control; promoter;
                                                                                                                                                                                                                                                                                                                                                                                          |: : :| : | 1: | 129 itprialttaeylayecgkhviviltdmssyadalrevsaareevpgrrgypgymytdla 308
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       357 rqlhnrqiyp-----pinvlpslsrlmksaigegmtrkdhsdvsnqlyanyaigkdvqa 410
                                                                                                                                                                                                                                                                                 ---NFGIHFMRVSSDLFPFASHAKYGYTLEFAQSHLEEVG------KLANK----- 118
                                                                                                                                                                                                                                                                                                                                                199 gednfalvfaamgvnm----etaqffkrdfeengsmervtlflnlandptieri 248
                                                                                                                                                                                                                                                                                                                                                                         119 YNHRLTMHPGQYTQIASPREVVV-----DSAIRDL-AYHDEILSRMKLNEQLNKDAV 169
                                                                                                                                                                                                                                                                                                                                                                                                                             170 LIIHLGGTFEGKKETLDRFRKNYQRLSDSVKARLVLENDDVSWSVQDLLPLCQELNIPLV 229
                                                                                                                                                                                                                                                                                                                                                                                                                                                       309 tiyeragriegrkgsitgi------piltmpnddithptpdltgyiteggiyid 356
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                230 LDWHHHNIVPGTLREGSLDLMPLIPTIRETWTRKGITQKQH-----YSESADPTAISG 282
                                                                                                                                                                                                                                         Gaps
                                                                                                                                                                                                                                                                  27 LNTILRSMKERVF-----CSRTCRITTIQRDGLESVKQLGTQNVLDLIKLVEWNH 76
                                                                                                                                                                                                                                         60; Mismatches 123; Indels 93;
                                                                                                                                                                                                         5.5%; Score 108.5; DB 21; Length 486; 20.0%; Pred. No. 0.059;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            Arabidopsis thaliana protein fragment SEQ ID NO: 20209.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   283 MKRRAHSDRVFDFPPCDPTMDLM-IEAKEKEQAVFELCRRYELQN 326
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               AAG18697 standard; Protein; 615 AA.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             99US-0121825.
99US-0123180.
99US-0123548.
99US-0125788.
99US-0126264.
                                                  990S-0160989.
990S-0161404.
990S-0161405.
990S-0161359.
                                                                                                                                              99US-0161920,
99US-0161992,
99US-0161993.
99US-0160814.
99US-0160815.
99US-0160980.
                                                                                                                     99US-0161360.
99US-0161361.
                                         99US-0160981.
                                                                                                                                                                                      99US-0162142
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       25-FEB-2000; 2000EP-0301439
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  17-OCT-2000 (first entry)
                                                                                                                                                                                                                                           Conservative
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 termination sequence.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          Arabidopsis thaliana.
                                                                                                                                                                                                                            Best Local Similarity
Matches 69; Conserv
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             25-FEB-1999;
05-MAR-1999;
09-MAR-1999;
23-MAR-1999;
25-MAR-1999;
                                                                                                                                                         28-OCT-1999;
28-OCT-1999;
29-OCT-1999;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    FP1033405-A2.
                                        22-OCT-1999;
22-OCT-1999;
25-OCT-1999;
                                                                                                                     26-0CT-1999;
26-0CT-1999;
                                                                               25-OCT-1999;
25-OCT-1999;
                                                                                                          26-OCT-1999;
                                                                                                                                                 28-OCT-1999;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            06-SEP-2000
             21-0CT-1999;
22-0CT-1999;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         AAG18697;
                                                                                                                                                                                                                 Query Match
                                                                                                                                                                                                                                                                                                                        17
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    AAG1.8697
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       RESULT
                                                                                                                                                                                                                                                                                                                                                 q
                                                                                                                                                                                                                                                                                                                                                                                                                                                         a
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  QY
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            qq
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         ò
                                                                                                                                                                                                                                                                                                                                                                                                    qq
                                                                                                                                                                                                                                                                                                                                                                                                                                ŏ
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     ò
                                                                                                                                                                                                                                                                                                                                                                            à
```

9 9 9 9 9 9 9 9 9 9 9 9 9 9 9 9 9 9 9	99; 99US-0147493. 99; 99US-0148711. 99; 99US-0148311. 99; 99US-0148684. 99; 99US-0149788. 99; 99US-0149788. 99; 99US-0149788. 99; 99US-0149723. 99; 99US-0149723. 99; 99US-0149723. 99; 99US-0149723. 99; 99US-015188. 99; 99US-015188.
19-JUL-199 19-JUL-199 19-JUL-199 19-JUL-199 20-JUL-199 20-JUL-199 21-JUL-199 21-JUL-199 22-JUL-199 22-JUL-199 22-JUL-199 23-JUL-199 23-JUL-199 23-JUL-199 23-JUL-199 23-JUL-199 23-JUL-199 23-JUL-199 23-JUL-199 23-JUL-199 23-JUL-199 23-JUL-199 23-JUL-199 23-JUL-199 24-JUL-199 27-JUL-199	09-AUG-1999 10-AUG-19999 11-
0 - sno	990S-0138094. 990S-0138094. 990S-0138640. 990S-0138452. 990S-0139452. 990S-0139454. 990S-0139456. 990S-0139456. 990S-0139456. 990S-0139457. 990S-0139460. 990S-0139461. 990S-0139461. 990S-0139461. 990S-0139463. 990S-0140991. 990S-0140991. 990S-0140991. 990S-0140991. 990S-0140991. 990S-0140991. 990S-0140991. 990S-0140991. 990S-0142920. 990S-0142920. 990S-014354.
APR-199	08-JUN-1999; 10-JUN-1999; 10-JUN-1999; 16-JUN-1999; 16-JUN-1999; 18-JU

```
29-MAR-1999;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                10-JUN-1
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      -JUN-1
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          18-JUN-1
21-JUN-1
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        18-JUN-
18-JUN-
18-JUN-
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            4-JUN-
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               8-JUN-
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      8-JUN-
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        3-JUN-
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            -NUL-8
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              03-JUN-
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  11;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   Protein identification; signal transduction pathway; metabolic pathway; hybridisation assay; genetic mapping; gene expression control; promoter;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            217 LLPLCQEL------TWTRKGI 265
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      162 EQLNKDAVLIIH----LGGTF-EGKKETLDRFRKNYQRLSDSVKARLVLENDDVSWSVQD 216
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                412 akisgddvvavnnmrligpvhteckksttgsfplkfdvlktkraarrdrvgetgtwglsr 471
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          48 TIQRDGLESVKQL-----GTQNVLDLIKLVEWNHNFGIHFMRVSSDLFPFASHAKYGYT 101
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  344 timegglrdlgglegdlvfgdagrkdvikflstnh-------iish---- 382
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 102 LEFAQSHLEEVGKLANKYNHRLTMHPGQYTQIASPREVVVDSAIRDLAYHDEILSRMKLN 161
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  77;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   5.4%; Score 107.5; DB 21; Length 615; 22.6%; Pred. No. 0.1;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  98; Indels
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           Arabidopsis thaliana protein fragment SEQ ID NO: 64427.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                34; Mismatches
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           266 TOKOHYSESADPTAISGMKRRAHSDRVFDF 295
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            AAG50810 standard; Protein; 615 AA.
9908 - 0159293 - 9908 - 0159294 - 9908 - 0159294 - 9908 - 0159329 - 9908 - 0159330 - 9908 - 0159331 - 9908 - 0160770 - 9908 - 0160981 - 9908 - 0160981 - 9908 - 0160981 - 9908 - 0161408 - 9908 - 0161408 - 9908 - 0161408 - 9908 - 0161408 - 9908 - 0161360 - 9908 - 0161360 - 9908 - 0161360 - 9908 - 0161360 - 9908 - 0161360 - 9908 - 0161360 - 9908 - 0161360 - 9908 - 0161360 - 9908 - 0161360 - 0908 - 0161360 - 0908 - 0161360 - 0908 - 0161360 - 0908 - 0161360 - 0908 - 0161360 - 0908 - 0161360 - 0908 - 0161360 - 0908 - 0161360 - 0088 - 0161360 - 0161360 - 0088 - 0161360 - 0088 - 0161360 - 0088 - 0161360 - 0088 - 0161360 - 0088 - 0161360 - 0161360 - 0161360 - 0161360 - 0161360 - 0161360 - 0161360 - 0161360 - 0161360 - 0161360 - 0161360 - 0161360 - 0161360 - 0161360 - 0161360 - 0161360 - 0161360 - 0161360 - 0161360 - 0161360 - 0161360 - 0161360 - 0161360 - 0161360 - 0161360 - 0161360 - 0161360 - 0161360 - 0161360 - 0161360 - 0161360 - 0161360 - 0161360 - 0161360 - 0161360 - 0161360 - 0161360 - 0161360 - 0161360 - 0161360 - 0161360 - 0161360 - 0161360 - 0161360 - 0161360 - 0161360 - 0161360 - 0161360 - 0161360 - 0161360 - 0161360 - 0161360 - 0161360 - 0161360 - 0161360 - 0161360 - 0161360 - 0161360 - 0161360 - 0161360 - 0161360 - 0161360 - 0161360 - 0161360 - 0161360 - 0161360 - 0161360 - 0161360 - 0161360 - 0161360 - 0161360 - 0161360 - 0161360 - 0161360 - 0161360 - 0161360 - 0161360 - 0161360 - 0161360 - 0161360 - 0161360 - 0161360 - 0161360 - 0161360 - 0161360 - 0161360 - 0161360 - 0161360 - 0161360 - 0161360 - 0161360 - 0161360 - 0161360 - 0161360 - 0161360 - 0161360 - 0161360 - 0161360 - 0161360 - 0161360 - 0161360 - 0161360 - 0161360 - 0161360 - 0161360 - 0161360 - 0161360 - 0161360 - 0161360 - 0161360 - 0161360 - 0161360 - 0161360 - 0161360 - 0161360 - 0161360 - 0161360 - 0161360 - 0161360 - 0161360 - 0161360 - 0161360 - 0161360 - 0161360 - 0161360 - 0161360 - 0161360 - 0161360 - 0161360 - 0161360 - 0161360 - 0161360 - 0161360 - 0161360 - 0161360 - 0161360 - 0161360 - 0161360 - 0161360 - 0161360 - 0161360 - 0161360
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  990S-0161361.
990S-0161920.
990S-0161992.
990S-0161993.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 18-OCT-2000 (first entry)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  61; Conservative
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  termination sequence.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            Arabidopsis thaliana.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             Similarity
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       EP1033405-A2.
                                                                                                                                                                           18 - OCT - 1999)
21 - OCT - 1999)
22 - OCT - 1999)
22 - OCT - 1999)
25 - OCT - 1999)
25 - OCT - 1999)
26 - OCT - 1999
26 - OCT - 1999
26 - OCT - 1999)
27 - OCT - 1999
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               06-SEP-2000.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    AAG50810;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          Query Match
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               Local
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  Matches
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     AAG50810
    qq
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         qq
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    В
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 ŏ
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            δ
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      δŏ
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          ōλ
```

```
990S-0132407.
990S-0132484.
990S-0132485.
990S-0132486.
99US-0132486.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                        990S-0137724.
990S-0138094.
990S-0138540.
990S-0138847.
990S-0139119.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 990S-0139452.
990S-0139453.
990S-0139492.
990S-0139454.
990S-0139455.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     990S-0139461.
99US-0139462.
99US-0139463.
99US-0139763.
99US-0139763.
                                                                                                                                                   990S-0130449.
990S-0130510.
990S-0130891.
990S-0131449.
990S-0132048.
                                                                                                                                                                                                                                                                                              990S-0134218.
990S-0134219.
990S-0134221.
990S-0134370.
990S-0134768.
                                                                                                                                                                                                                                                                                                                                                                   99US-0135124.
99US-0135353.
99US-0135629.
                                                                                                                                                                                                                                                                                                                                                                                                     99US-0136021.
99US-0136392.
99US-0136782.
                                                                                                                                                                                                                                                                                                                                                                                                                                      99US-0137222.
99US-0137528.
99US-0137502.
                                                                                                                  99US-0128714.
99US-0129845.
99US-0130077.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             99US-0139456.
99US-0139457.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    99US-0139458.
99US-0139459.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          99US-0139460.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             99US-0140695.
                                                                                                                                                                                                                                                                        99US-0132863.
99US-0134256.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          990S-0139899.
990S-0140353.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              99US-0141287.
25-FEB-2000; 2000EP-0301439.
                       99US-0121825
                                                                    99US-0126264.
                                                                                            99US-0127462.
99US-0128234.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    99US-0140991
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                99US-0142055
99US-0142390
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       99US-0142154
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  99US-01
                                                                                                                                                                                                                                                                                              14-MAY-1999;
14-MAY-1999;
14-MAY-1999;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          17-JUN-1999;
18-JUN-1999;
18-JUN-1999;
                                                                                                                                                                                        28-APR-1999;
30-APR-1999;
30-APR-1999;
                                                                                                                                                                                                                                                                                                                                                                                                                                         1999;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                         07-JUN-1999;
08-JUN-1999;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             1999;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             18-JUN-1999;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                :666
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             666
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  1999;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          22-JUN-1999;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             24-JUN-1999;
28-JUN-1999;
                                                                                                                                                                  23-APR-1999;
23-APR-1999;
                                                                                                                                                                                                                                                                        07-MAY-1999;
11-MAY-1999;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    6-JUN-1999;
6-JUN-1999;
                                                                                                                   08-APR-1999;
16-APR-1999;
                                                                                                                                           19-APR-1999;
21-APR-1999;
                                                                                                                                                                                                                                      05-MAY-1999;
06-MAY-1999;
```

us-09-724-296-4.rag

```
990S - 0142920
990S - 0142920
990S - 0143542
990S - 0144065
990S - 0144085
990S - 0144085
990S - 0144331
990S - 0144332
990S - 0144334
990S - 0144334
990S - 0144336
990S - 0144632
990S - 0144632
990S - 0144632
990S - 014684
990S - 014684
990S - 014686
                                                                                                                                        99US-0145089.
99US-0145192.
99US-0145145.
                                                                                                                                                                                   990S-0145918.
990S-0145919.
990S-0145951.
990S-0146388.
990S-0146388.
                                                                                                                                                                                                                                99US-0147204.
99US-0147302.
99US-0147192.
99US-0147260.
                                                                                                                                                                                                                                                                                   99US-0148171.
99US-0148319.
99US-0148341.
                                                                                                                                                                                                                                                                                                     99US-0148565.
99US-0148684.
99US-0149368.
                                                                                                                                                                                                                                                                                                                              99US-0149426.
99US-0149722.
99US-0149723.
                                                                                                                                                                                                                                                                                                                                                                           99US-0150884.
99US-0151065.
99US-0151066.
                                                                                                                                                                                                                                                                                                                                                                                                                                                99US-0154039.
99US-0154779.
99US-0155139.
99US-0155486.
                                                                                                                                                           99US-0145218.
99US-0145224.
                                                                                                                                                                       99US-0145276.
99US-0145913.
                                                                                                                                                                                                                                                          99US-0147303.
99US-0147416.
                                                                                                                                                                                                                                                                                                                                                 99US-0149929.
99US-0149902.
                                                                                                                                                                                                                                                                                                                                                                                                          99US-0151438.
99US-0151930.
                                                                                                                                                                                                                                                                                                                                                                                                                             99US-0153070.
                                                                                                                                                                                                                                                                             90S-0147935
                                                                                                                                                                                                                                                                                                                        99US-0149175
                                                                                                                                                                                                                                                                                                                                                              99US-0149930
                                                                                                                                                                                                                                                                                                                                                                     99US-0150566
                                                                                                                                                                                                                                                                                                                                                                                              39US-0151080
                                                                                                                                                                                                                                                                                                                                                                                                                       9US-0152363
                                                                                                                                                                                                                                                                      99US-0147493
                                                                                                                                                                                                                                                                                                                                                                                                                                          9US-0154018
                13-JUL-1999;
14-JUL-1999;
15-JUL-1999;
16-JUL-1999;
16-JUL-1999;
19-JUL-1999;
                                                                                                                                                                                                                                                                                                                                                                                                                                                     -SEP-1999;
-SEP-1999;
-SEP-1999;
                                                                   999;
                                                                                            6666
                                                                                                                                                                 999;
                                                                                                                                                                                                                         :6661
                                                                                                                                                                             999;
                                                                                                                                                                                                                                                                                                                                           :666
                                                                                                                                                                                                  666
                                                                                                                                                                                                                                                                                                                                                 20-AUG-19
23-AUG-19
                                                                                                                                                                                                                                                                                                                                                                                             27-AUG-19
30-AUG-19
31-AUG-19
01-SEP-19
                                                                                                                                                                                                                                                                                                     13-AUG-19
13-AUG-19
16-AUG-19
                                                                                                                                                                                                                                                                                                                                    20-AUG-19
20-AUG-19
                                                                                                                                                                                                       02-AUG-1
02-AUG-1
                                                                                                                                                                                                                                                               06-AUG-1
09-AUG-1
                                                                                                                                                                                                                    ·AUG-1
                                                                                                                                                                                                                                      04-AUG-1
                                                                                                                                                                                                                                                                                                                                                                                 27 - AUG - 1
27 - AUG - 1
                                                                                                                                                                                                                                                                                                                                                                                                                       SEP-1
                                                                                                                                              -JUL-1
                                                                                                                                                                                                                                                   05-AUG-1
                                                                                                                                                                                                                                                                             AUG-1
                                                                                                                                                                                                                                                                                                                        AUG-1
                                                                                                                                  22-JUL-1
                                                                                                                                                                                                                                                                                                                                                               -AUG-1
                                                                                                                                                                                                                                                                                                                                                                    5-AUG-1
                                                                                                                                                                                                                                                                                                                                                                           AUG-1
                                                                                                                                                                                                                                                                                                                                                                                                                              10-SEP-1
                                                                                                                                                                                                                                                                                                                                                                                                                                    SEP-1
                                                                                                                21-JUL-
                                                                                                                                                    23-JUL-
                                                                                                                                                                 23-JUL-
                                                                                                                                                                                                                                                                                         - AUG-
                                                                                                                                                                                                                                                                                                                                                                                                                                                 SEP-
                                                                                                                                                           23-JUL-
```

```
217 LLPLCQEL-----INPLVLDWHHNIVPGTLREGSL--DLMPLIPTIRE--TWTRKGI 265
                                                                                                                                                                                                                                                                                                                                                                                                                          || :|| :|| || 344 timegglrdlgglegdlvfgdagrkdvikflstnh-------iish----- 382
                                                                                                                                                                                                                                                                                                                                                                                                                                                  102 LEFAQSHLEEVGKLANKYNHRLTMHPGQYTQIASPREVVVDSAIRDLAYHDEILSRMKLN 161
                                                                                                                                                                                                                                                                                                                                                                                                                                                              EQLANKDAVLIIH----LGGTF-EGKKETLDRFRKNYQRLSDSVKARLVLENDDVSWSVQD 216
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   472 fypiveelveklskghlp-kqdypcmnepkptfysgslspsaspvlphsrrtptwarrhl 530
                                                                                                                                                                                                                                                                                                                                                                                                             TIQRDGLESVKQL-----GTQNVLDLIKLVEWNHNFGIHFMRVSSDLFPFASHAKYGYT 101
                                                                                                                                                                                                                                                                                                                                                                        DB 21; Length 615;
                                                                                                                                                                                                                                                                                                                                                                                           Indels
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                Arabidopsis thaliana protein fragment SEQ ID NO: 20208.
                                                                                                                                                                                                                                                                                                                                                                      5.4%; Score 107.5; [2.6%; Pred. No. 0.1;
                                                                                                                                                                                                                                                                                                                                                                                          34; Mismatches
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    266 TQKQHYSESADPTAISGMKRRAHSDRVFDF 295
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              AAG18696 standard; Protein; 633 AA
                                                        990S-0158029.
990S-0158232.
990S-0158369.
990S-0159293.
990S-0159294.
990S-0155659.
990S-0156458.
990S-0156596.
990S-0157117.
990S-0157753.
                                                                                                                                                                                                                                990S-0160980.
99US-0160981.
99US-0160989.
                                                                                                                                                                                                                                                                                                                                 99US-0161992.
99US-0161993.
99US-0162142.
                                                                                                                                                               99US-0159584.
                                                                                                                                                                                                      99US-0160770.
                                                                                                         99US-0159295.
                                                                                                                            99US-0159330
                                                                                                                                                                                              99US-0160768
                                                                                                                                                                                                                                                                                           99US-0161359
                                                                                                                                                                                                                                                                                                                                                                                22.6%;
                                                                                                                                    99US-0159331
99US-0159637
                                                                                                                                                        99US-0159638
                                                                                                                                                                                   99US-0160767
                                                                                                                                                                                                                          99US-0160815
                                                                                                                                                                                                                                                                        99US-0161405
                                                                                                                                                                                                                                                                                  99US-0161406
                                                                                                                                                                                                                                                                                                      99US-0161360
                                                                                                                                                                                                                                                                                                               99US-0161361
                                                                                                                                                                                                                                                                                                                        99US-0161920
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           17-OCT-2000 (first entry)
                                                                                                                                                                                                                                                                                                                                                                                Local Similarity 22.6 es 61; Conservative
                                                                                                                                                                                                                                                                99US-01
24-SEP-1999;
28-SEP-1999;
04-0CT-1999;
05-0CT-1999;
06-0CT-1999;
07-0CT-1999;
13-0CT-1999;
13-0CT-1999;
13-0CT-1999;
14-0CT-1999;
                                                                                                                                                                                                                        21-OCT-1999;
22-OCT-1999;
                                                                                                                                                                                                                                                                                          26-0CT-1999;
26-0CT-1999;
26-0CT-1999;
28-0CT-1999;
28-0CT-1999;
                                                                                                                                                       .4-OCT-1999;
                                                                                                                                                               8-OCT-1999;
                                                                                                                                                                                   1-0CT-1999;
                                                                                                                                                                                              21-OCT-1999;
                                                                                                                                                                                                      21-OCT-1999;
21-OCT-1999;
                                                                                                                                                                                                                                            22-OCT-1999;
                                                                                                                                                                                                                                                    22-0CT-1999;
25-0CT-1999;
                                                                                                                                                                                                                                                                       25-OCT-1999;
25-OCT-1999;
                                                                                                                                                                                                                                                                                                                                                                       Query Match
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         AAG18696;
                                                                                                                                    14-0CT-1
14-0CT-1
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        162
                                                                                                                                                                                                                                                                                                                                                                                                             48
                                                                                                                                                                                                                                                                                                                                                                                          Matches
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              AAG18696
  qa
                                                                                                                                                                                                                                                                                                                                                                                                              ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                   Ωy
                                                                                                                                                                                                                                                                                                                                                                                                                                                                    Db
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          qq
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              Ω
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                Op
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    Qγ
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     Dp
```

```
PR
Protein identification; signal transduction pathway; metabolic pathway; hybridisation assay; genetic mapping; gene expression control; promoter;
                                                                                                                            990xS-0121825
990xS-0123180
990xS-0125788
990xS-0125788
990xS-0126785
990xS-01274625
990xS-01287144
990xS-0130047
990xS-0130040
990xS-0131449
990xS-0132486
990xS-0132486
990xS-0132486
990xS-0132486
990xS-0132486
990xS-0132487
990xS-0134748
990xS-0139455
990xS-0139455
990xS-0139455
990xS-0139455
990xS-0139455
990xS-0139455
990xS-0139455
990xS-0139455
990xS-0139455
990xS-0139453
990xS-0139453
990xS-0139453
990xS-0139463
990xS-0139463
990xS-0139463
                                                                                                          25-FEB-2000; 2000EP-0301439
                        termination sequence.
                                            Arabidopsis thaliana
                                                                                                                            25 - FEB - 1999

05 - MAR - 1999

23 - MAR - 1999

25 - MAR - 1999

10 - APR - 1999

10 - APR - 1999

10 - APR - 1999

11 - APR - 1999

12 - APR - 1999

13 - APR - 1999

14 - MAY - 1999

15 - MAY - 1999

16 - MAY - 1999

17 - MAY - 1999

18 - MAY - 1999

19 - MAY - 1999

19 - MAY - 1999

19 - MAY - 1999

10 - JUN - 1999

11 - JUN - 1999

12 - MAY - 1999

13 - MAY - 1999

14 - MAY - 1999

16 - JUN - 1999

17 - JUN - 1999

18 - JUN - 1999

19 - JUN - 1999

10 - JUN - 1999

11 - JUN - 1999
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                18 - JUN - 1999

18 - JUN - 1999

18 - JUN - 1999

18 - JUN - 1999

18 - JUN - 1999

18 - JUN - 1999

18 - JUN - 1999

18 - JUN - 1999

18 - JUN - 1999

22 - JUN - 1999

23 - JUN - 1999

23 - JUN - 1999

23 - JUN - 1999
                                                               EP1033405-A2.
                                                                                    06-SEP-2000.
```

905 - 01400354 905 - 01406955 905 - 01406923 905 - 0141287 905 - 0141287 905 - 01421845 905 - 01421845 905 - 0142056 905 - 0142090 905 - 0142090 905 - 0142090	990S-0144065 990S-0144086 990S-0144086 990S-014432 990S-014433 990S-014433 990S-014433 990S-014433 990S-014433 990S-014434 990S-014484 990S-014484 990S-014508 990S-014508 990S-014508 990S-014508 990S-014508 990S-0145192 990S-0145192 990S-0145193 990S-0145193 990S-014518 990S-014518 990S-014518 990S-014518 990S-014518 990S-014518	90.5 0147302 90.5 0147303 90.5 0147303 90.5 0147303 90.5 0147493 90.5 0147493 90.5 0148311 90.5 0148311 90.5 0148368 90.5 0148368 90.5 0148368 90.5 0149368 90.5 0149303 90.5 0149903 90.5 0149903 90.5 0149903 90.5 0149903 90.5 0149903 90.5 0149903 90.5 015088 90.5 0149903 90.5 015088 90.5 015088 90.5 015088 90.5 0151088
23 - JUN - 1999 24 - JUN - 1999 29 - JUN - 1999 30 - JUN - 1999 01 - JUL - 1999 02 - JUL - 1999 08 - JUL - 1999 09 - JUL - 1999 13 - JUL - 1999	PR 14-JUL-1999; PR 16-JUL-1999; PR 16-JUL-1999; PR 19-JUL-1999; PR 19-JUL-1999; PR 19-JUL-1999; PR 19-JUL-1999; PR 20-JUL-1999; PR 20-JUL-1999; PR 20-JUL-1999; PR 20-JUL-1999; PR 21-JUL-1999; PR 22-JUL-1999; PR 22-JUL-1999; PR 22-JUL-1999; PR 22-JUL-1999; PR 22-JUL-1999; PR 22-JUL-1999; PR 23-JUL-1999; PR 23-JUL-1999	R

```
AAG50809 standard; Protein; 633 AA.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     99US-0138847.
99US-0139119.
99US-0139452.
                                                                                                                                                                                                                                                            25-FEB-2000; 2000EP-0301439
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            99US-0139453
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         99US-0139492
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       99US-0139454
                                                                      18-OCT-2000 (first entry)
                                                                                                                                                                             Arabidopsis thallana.
                                                                                                                                                                                                                                                                                                                          23-MAR-1999;
25-MAR-1999;
29-MAR-1999;
01-APR-1999;
                                                                                                                                                                                                       EP1033405-A2
                                                                                                                                                                                                                                                                                                                                                                                                         16-APR-1999;
19-APR-1999;
                                                                                                                                                                                                                                                                                                                                                                                                                                   21-APR-1999;
23-APR-1999;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                         28-APR-1999;
30-APR-1999;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          :666
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                :666
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            :666
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               .666
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     666
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              :666
                                                                                                                                                                                                                                                                                                05-MAR-1999;
09-MAR-1999;
                                                                                                                                                                                                                                                                                                                                                                              06-APR-1999;
08-APR-1999;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       :666
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    , 666
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              :666
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            -JUN-1999;
                                                                                                                                                                                                                                  06-SEP-2000
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    30-APR-1999;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               -MAY-1999;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       .0-JUN-1999;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            6-JUN-1999;
                                             AAG50809;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             -MAY-1
-MAY-1
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             MAY-1
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    MAY-1
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               MAY - 3
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          - MAY - 1
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       - MAY - J
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     - MAY - 1
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             MAY - 1
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              -JUN-1
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 14-JUN-1
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     -JUN-
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            -JUN-
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      -NOC-01
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             11;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            Ouery Match 5.4%; Score 107.5; DB 21; Length 633; Best Local Similarity 22.6%; Pred. No. 0.11; Matches 61; Conservative 34; Mismatches 98; Indels 77; Gaps
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     LLPLCQEL-----NIPLVLDWHHHNIVPGTLREGSL--DLMPLIPTIRE--TWTRKGI 265
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            || || :| || || || 362 timegglrdlgglegdlvfgdagrkdvikflstnh-------iish----- 400
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   102 LEFAQSHLEEVGKLANKYNHRLIMHPGQYTQIASPREVVVDSAIRDLAYHDEILSRMKLN 161
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         EQLNKDAVLIIH----LGGTF-EGKKETLDRFRKNYQRLSDSVKARLVLENDDVSWSVQD 216
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               48 TIQRDGLESVKQL-----GTQNVLDLIKLVEWNHNFGIHFMRVSSDLFPFASHAKYGYT 101
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               266 TQKQHYSESADPTAISGMKRRAHSDRVFDF 295
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               9905-0151930.
9905-0152363.
9905-01533070.
9905-0153758.
9905-0154018.
9905-0154779.
9905-0155139.
9905-0155486.
9905-0155486.
                                                                                                                                                                                                                                                                990S-0159295
990S-0159329
990S-0159330
990S-0159331
990S-0159637
990S-0159584
                                                                                                                                                                                    990S-0157865.
990S-0158029.
990S-0158232.
990S-0158369.
990S-0159293.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                    99US-0160989.
99US-0161404.
99US-0161405.
99US-0161406.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          99US-0161992.
99US-0161993.
99US-0162142.
                                                                                                                                                                                                                                                                                                                                                                                                                  99US-0160814
                                                                                                                                                             99US-0157117
                                                                                                                                                                         99US-0157753
                                                                                                                                                                                                                                                                                                                                                                            99US-0160767
                                                                                                                                                                                                                                                                                                                                                                                       99US-0160768
                                                                                                                                                                                                                                                                                                                                                                                                      99US-0160770
                                                                                                                                                                                                                                                                                                                                                                                                                                           99US-0160980
99US-0160981
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     99US-0161360
99US-0161361
01-SEP-1999;
07-SEP-1999;
13-SEP-1999;
15-SEP-1999;
16-SEP-1999;
20-SEP-1999;
23-SEP-1999;
24-SEP-1999;
24-SEP-1999;
24-SEP-1999;
24-SEP-1999;
                                                                                                                                            29 - SEP - 1999

06 - OCT - 1999

07 - OCT - 1999

12 - OCT - 1999

13 - OCT - 1999

14 - OCT - 1999

14 - OCT - 1999

14 - OCT - 1999

16 - OCT - 1999

17 - OCT - 1999

18 - OCT - 1999

19 - OCT - 1999

21 - OCT - 1999

21 - OCT - 1999

22 - OCT - 1999

25 - OCT - 1999

25 - OCT - 1999

26 - OCT - 1999

27 - OCT - 1999

28 - OCT - 1999

29 - OCT - 1999

20 - OCT - 1999

20 - OCT - 1999

21 - OCT - 1999

22 - OCT - 1999

22 - OCT - 1999

23 - OCT - 1999

25 - OCT - 1999

26 - OCT - 1999

27 - OCT - 1999

27 - OCT - 1999

28 - OCT - 1999

27 - OCT - 1999

28 - OCT - 1999

27 - OCT - 1999

28 - OCT - 1999

27 - OCT - 1999

28 - OCT - 1999

27 - OCT - 1999

28 - OCT - 1999

29 - OCT - 1999

20 - OCT - 1999

21 - OCT - 1999

22 - OCT - 1999

23 - OCT - 1999

24 - OCT - 1999

25 - OCT - 1999

26 - OCT - 1999

27 - OCT - 1999

27 - OCT - 1999

28 - OCT - 1999

27 - OCT - 1999

28 - OCT - 1999

27 - OCT - 1999

28 - OCT - 1999

29 - OCT - 1999
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              26-0CT-1999;
28-0CT-1999;
28-0CT-1999;
28-0CT-1999;
29-0CT-1999;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                401
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         162
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            217
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        549
  q
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             qq
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               QQ
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       οy
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                à
```

RESULT

```
Protein identification; signal transduction pathway; metabolic pathway; hybridisation assay; genetic mapping; gene expression control; promoter; termination sequence.
                                       Arabidopsis thaliana protein fragment SEQ ID NO: 64426
                                                                                                                    9905-0121825.
9905-0123180.
9905-0125788.
9905-0125788.
9905-0126785.
9905-0126785.
9905-0126785.
9905-013077.
9905-0130670.
9905-0130670.
9905-01308945.
                                                                                                                                                                                                                                                                                                                  990S-01367822
990S-0137222
990S-0137528
990S-0137502
990S-0138094
                                                                                                                                                                                                                    990S-0132484.
990S-0132485.
990S-0132486.
990S-0132487.
990S-0134256.
990S-0134218.
990S-0134219.
990S-0134219.
990S-0134219.
                                                                                                                                                                                                                                                                                            990S-0135353.
990S-0135629.
990S-0136021.
990S-0136392.
                                                                                                                                                                                                                                                                                                                                                                                          990S-0139455.
990S-0139456.
990S-0139457.
990S-0139458.
                                                                                                                                                                                                                                                                                                                                                                                                     18-JUN-1999;
18-JUN-1999;
18-JUN-1999;
```

UUS-0139460 UUS-0139461 UUS-0139462 UUS-0139462 UUS-0139763 UUS-0139817 UUS-0139817 UUS-0140353 UUS-0140353 UUS-01408253 UUS-01408253 UUS-01408253 UUS-01408253	00S-0141287 00S-0141842 00S-0142803 00S-0142803 00S-0142803 00S-0142803 00S-0143842 00S-0144005 00S-0144005 00S-0144005 00S-0144005 00S-0144005 00S-0144005 00S-0144005 00S-0144005 00S-0144333	905-0144334 905-0144335 905-0144335 905-0144864 905-0144814 905-0145086 905-0145086 905-0145089 905-0145089 905-0145089 905-014518 905-014518	9905-0145951. 9905-0146386. 9905-0146389. 9905-0147038. 9905-0147304. 9905-0147304. 9905-0147305. 9905-0147305. 9905-0147416. 9905-0147416. 9905-0147319. 9905-0148319. 9905-0148664. 9905-0148319. 9905-0148319. 9905-0148319. 9905-0148319. 9905-0148319. 9905-0148319. 9905-0149368.
- JUN - 1999 - JUN - 1999		9-JUL-199 9-JUL-199 10-JUL-1	AUG-1999
7	9	2 2 2 2 2 2 2 2 2 2 2 2 2 2 2 2 2 2 2 2	

```
162 EQLNKDAVLIIH----LGGTF-EGKKETLDRFRKNYQRLSDSVKARLVLENDDVSWSVQD 216
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        Query Match 5.4%; Score 107.5; DB 21; Length 633; Best Local Similarity 22.6%; Pred. No. 0.11; Matches 61; Conservative 34; Mismatches 98; Indels 77; Gaps
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               102 LEFAQSHLEEVGKLANKYNHRLTMHPGQYTQIASPREVVVDSAIRDLAYHDEILSRMKLN 161
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             217 LLPLCQEL-----NIPLVLDWHHHNIVPGTLREGSL--DLMPLIPTIRE--TWTRKGI 265
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     48 TIQRDGLESVKQL-----GTQNVLDL1KLVEWNHNFGIHFMRVSSDLFPFASHAKYGYT 101
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           362 timegglrdlgglegdlvfgdagrkdvikflstnh-------iish----- 400
                                                                           990x - 0151303
990x - 0151303
990x - 0151438
990x - 0152363
990x - 0153070
990x - 0154018
990x - 0154018
990x - 0154018
990x - 0154739
990x - 0156559
990x - 0156559
990x - 015659
990x - 015659
990x - 0159293
                                                                                                                                                                                                                                                                                                                                                                                                990S-0160741
990S-0160767
990S-0160768
990S-0160710
990S-0160815
990S-0160981
990S-0160989
990S-0161404
990S-0161406
990S-0161406
990S-0161406
990S-0161406
99US-0149902.
99US-0149930.
99US-0150566.
99US-0150884.
99US-0151065.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   990S-0161920.
990S-0161992.
990S-0161993.
990S-0162142.
                                                     99US-0151066.
                                                                                                                                                                20 - SEP - 1999;
23 - SEP - 1999;
24 - SEP - 1999;
26 - SEP - 1999;
26 - SEP - 1999;
27 - SEP - 1999;
28 - SEP - 1999;
27 - OCT - 1999;
28 - OCT - 1999;
27 - OCT - 1999;
27 - OCT - 1999;
27 - OCT - 1999;
28 - OCT - 1999;
27 - OCT - 1999;
28 - OCT - 1999;
27 - OCT - 1999;
27 - OCT - 1999;
27 - OCT - 1999;
28 - OCT - 1999;
27 - OCT - 1999;
23-AUG-1999;
25-AUG-1999;
26-AUG-1999;
27-AUG-1999;
27-AUG-1999;
30-AUG-1999;
31-AUG-1999;
31-AUG-1999;
31-AUG-1999;
31-AUG-1999;
31-AUG-1999;
31-AUG-1999;
31-AUG-1999;
                                                                                                                                 13-SEP-1999;
15-SEP-1999;
16-SEP-1999;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         Oy
    qq
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 ŏ
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   qq
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  qq
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        ŏ
```

```
990S-0139459.
990S-0139460.
990S-0139461.
990S-0139462.
990S-0139463.
                         990S-0139454.
990S-0139455.
990S-0139456.
990S-0139457.
990S-0139458.
                                                                                                                                                                                                                                                                                                                                                                         990S-0144884.
990S-0144814.
990S-0145086.
990S-0145088.
990S-0145085.
                                                                                                                                                                                                                                                                                                                                                                                                                            990S-0145089.
990S-0145192.
990S-0145145.
990S-0145218.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                       99US-0145276.
99US-0145913.
99US-0145918.
99US-0145919.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            990S-0147416.
990S-0147493.
990S-0147935.
990S-0148171.
990S-0148319.
99US-0139452.
99US-0139453.
99US-0139492.
                                                                                                                      99US-0139763.
99US-0139817.
99US-0139899.
                                                                                                                                                                                                                                                                                                       99US-0144325.
99US-0144331.
99US-0144332.
                                                                                                                                                                                                                                                                                                                                                 99US-0144335.
99US-0144352.
99US-0144632.
                                                                                                                                                99US-0140353.
                                                                                                                                                                 99US-0140695.
                                                                                                                                                                                  99US-0140991
                                                                                                                                                                                                                            99US-0142390.
99US-0142803.
                                                                                                                                                                                                                                                                     99US-0143624
99US-0144005
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          99US-0145951.
99US-0146386.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     99US-0147192.
99US-0147260.
                                                                                                                                                                                                            99US-0142154
99US-0142055
                                                                                                                                                                                                                                             99US-0142920
99US-0142977
                                                                                                                                                                                                                                                              99US-0143542
                                                                                                                                                                                                                                                                                       99US-0144085
99US-0144086
                                                                                                                                                                                                                                                                                                                                 99US-0144333
99US-0144334
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           99US-0146388
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  99US-0146389
99US-0147038
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    99US-0147204
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             99US-0147302
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       99US-0147303
                                                                                                                                                                                          99US-0141287
                                                                                                                                                                                                   99us-01
                                                                                                                                                                                                                          06 - JUL - 1999;
08 - JUL - 1999;
08 - JUL - 1999;
12 - JUL - 1999;
13 - JUL - 1999;
15 - JUL - 1999;
16 - JUL - 1999;
16 - JUL - 1999;
16 - JUN - 1999;
16 - JUN - 1999;
18 - JUN - 1999;
23 - JUN - 1999;
                                                                                                                                                               24 - JUN - 1999;
28 - JUN - 1999;
29 - JUN - 1999;
30 - JUN - 1999;
01 - JUL - 1999;
02 - JUL - 1999;
                                                                                                                                                                                                                                                                                                       19-JUL-1999;
19-JUL-1999;
19-JUL-1999;
19-JUL-1999;
19-JUL-1999;
20-JUL-1999;
                                                                                                                                                                                                                                                                                                                                                                  20-JUL-1999;
20-JUL-1999;
21-JUL-1999;
21-JUL-1999;
21-JUL-1999;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 02-AUG-1999)
03-AUG-1999)
04-AUG-1999)
05-AUG-1999)
05-AUG-1999)
06-AUG-1999)
09-AUG-1999)
09-AUG-1999)
11-AUG-1999)
                                                                                                                                                                                                                                                                                                                                                                                                                             22-JUL-1999;
22-JUL-1999;
23-JUL-1999;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         27-JUL-1999;
27-JUL-1999;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 02-AUG-1999;
02-AUG-1999;
                                                                                                                                                                                                                                                                                                                                                                                                             22-JUL-1999;
22-JUL-1999;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                        26-JUL-1999;
27-JUL-1999;
                                                                                                                                                                                                                                                                                                                                                                                                                                                       23-JUL-1999;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                23-JUL-1999;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           28-JUL-1999;
 ********************************
                                                                                                                                                     Protein identification; signal transduction pathway; metabolic pathway; hybridisation assay; genetic mapping; gene expression control; promoter; termination sequence.
Arabidopsis thaliana protein fragment SEQ ID NO: 20207.
                         266 TQKQHYSESADPTAISGMKRRAHSDRVFDF 295
                                   sddgyfsdsvlgrassgfkrk--gqrifvf 576
                                                                                   AAG18695 standard; Protein; 662 AA
                                                                                                                                                                                                                                                                    990S - 0123180
990S - 0123548
990S - 0126264
990S - 0126264
990S - 0126264
990S - 0126274
990S - 01282714
990S - 0130077
990S - 0130610
990S - 0130610
990S - 0130610
990S - 0131469
990S - 0132467
990S - 0132487
990S - 0132486
990S - 0132486
990S - 0132487
990S - 0132486
990S - 0132486
990S - 0132486
990S - 0134218
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         990S-0136021
99US-0136392
99US-0136782
99US-0137228
99US-0137228
99US-0137724
99US-0138094
99US-0138094
99US-0138094
99US-0138094
                                                                                                                                                                                                                                           25-FEB-2000; 2000EP-0301439
                                                                                                                                                                                                                                                             99US-0121825
                                                                                                                      17-OCT-2000 (first entry)
                                                                                                                                                                                         Arabidopsis thaliana
                                                                                                                                                                                                                                                           25-FEB-1999;
05-MAR-1999;
23-MAR-1999;
25-MAR-1999;
29-MAR-1999;
06-APR-1999;
06-APR-1999;
19-APR-1999;
19-APR-1999;
23-APR-1999;
23-APR-1999;
30-APR-1999;
30-APR-1999;
30-APR-1999;
30-APR-1999;
66-MAY-1999;
06-MAY-1999;
                                                                                                                                                                                                                                                                                                                                                                                                                                             07-MAY-1999;
11-MAY-1999;
14-MAY-1999;
14-MAY-1999;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                               14-MAY-1999;
14-MAY-1999;
18-MAY-1999;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       19-MAY-1999;
20-MAY-1999;
21-MAY-1999;
24-MAY-1999;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         25-MAY-1999;
27-MAY-1999;
28-MAY-1999;
01-JUN-1999;
04-JUN-1999;
07-JUN-1999;
                                                                                                                                                                                                          EP1033405-A2
                                                                                                                                                                                                                           06-SEP-2000
                                                                                                     AAG18695;
                                                                  RESULT 15
                                         549
                                                                           AAG18695
        q
                                                                                             ò
                                         g
```

```
9905-0148341.
9905-0148565.
9905-0148565.
9905-0149175.
9905-0149726.
9905-0149723.
9905-0149929.
9905-0149929.
9905-0150884.
9905-0150884.
9905-0151080.
9905-0151080.
9905-0151080.
9905-0151330.
9905-0151330.
9905-0151330.
9905-0151330.
9905-0151330.
9905-0151330.
9905-0151330.
9905-0151330.
9905-0151330.
9905-0151330.
9905-0151330.
9905-0151330.
9905-0159330.
9905-0159330.
9905-0159330.
9905-0159330.
9905-0159330.
9905-0159330.
9905-0159330.
9905-0159330.
9905-0159330.
9905-0159330.
9905-0159330.
9905-0159331.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            990S-0161361.
990S-0161920.
990S-0161992.
990S-0161993.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             99US-0160989.
99US-0161404.
99US-0161405.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              99US-0161406.
99US-0161359.
99US-0161360.
                                                                                                                                                                                                                                                                                                                                                                                                                                     23 - SEP - 1999
24 - SEP - 1999
28 - SEP - 1999
04 - OCT - 1999
06 - OCT - 1999
06 - OCT - 1999
08 - OCT - 1999
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         21-007-1999
21-007-1999
22-007-1999
22-007-1999
25-007-1999
25-007-1999
26-007-1999
26-007-1999
26-007-1999
28-007-1999
28-007-1999
28-007-1999
12-AUG-1999;
13-AUG-1999;
16-AUG-1999;
17-AUG-1999;
18-AUG-1999;
20-AUG-1999;
20-AUG-1999;
23-AUG-1999;
23-AUG-1999;
23-AUG-1999;
27-AUG-1999;
30-AUG-1999;
31-AUG-1999;
31-AU
                                                                                                                                                                                                                                                                                                                                                                                       16-SEP-1999;
20-SEP-1999;
22-SEP-1999;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             14-OCT-1999;
14-OCT-1999;
14-OCT-1999;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           21-0CT-1999;
21-0CT-1999;
21-0CT-1999;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            13-OCT-1999;
13-OCT-1999;
                                                                                                                                                                                                                                                                                                                                                                         15-SEP-1999;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              12-OCT-1999;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               13-OCT-1999;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            14-OCT-1999;
14-OCT-1999;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               18-OCT-1999;
```

```
Ouery Match

5.4%; Score 107.5; DB 21; Length 662;
Best Local Similarity 22.6%; Pred. No. 0.12;
Matches 61; Conservative 34; Mismatches 98; Indels 77;
```

48 TIQRDGLESVKQL-----GTQNVLDLIKLVEWNHNFGIHFMRVSSDLFPFASHAKYGYT 101

рp δλ

11;

Qy	102 LEFAQSHLEEVGKLANKYNHRLTMHPGQYTQIASPREVVVDSAIRDLAYHDEILSRMKLN 161	
Db	430eskl	
QY	162 EQLNKDAVLIIHLGGTF-EGKKETLDRFRKNYQRLSDSVKARLVLENDDVSWSVQD 216	
qq	459 akisgddyvavnnmrligpvhteckkstigsfplkfdvlktkraarrdrygetgtwgisr 518	
Qy	217 LLPLCQELNIPLVLDWHHHNIVPGTLREGSLDLMPLIPTIRETWTRKGI 265	
QQ	519 fypiveelveklskgilp-kgdypcmnepkptfysgslspsaspvlphsrrtptwarrhl 577	
QY	266 TOKOHYSESADPTAISGMKRRAHSDRVFDF 295	
qq	578 sddgyfsdsvlgrassgfkrkggrifvf 605	

Search completed: January 15, 2002, 13:50:00 Job time: 289 sec

```
GenCore version 4.5
Copyright (c) 1993 - 2000 Compugen Ltd.
```

OM protein - protein search, using sw model

January 15, 2002, 13:50:15 Run on:

; Search time 120.01 Seconds
(without alignments)
181.464 Million cell updates/sec

US-09-724-296-39 1527 Title:

1 QLGLVCLTVGPEVRFRTVTL......VEAKGKEEAIAALRLMAPFK 294 Perfect score:

Sequence:

Scoring table:

522463 seqs, 74073290 residues Searched:

BLOSUM62 Gapop 10.0 , Gapext 0.5

Total number of hits satisfying chosen parameters:

Minimum DB seq length: 0 Maximum DB seq length: 200000000

Post-processing: Minimum Match 0% Maximum Match 100% Listing first 45 summaries

Database

/SIDS8/gcgdata/geneseq/geneseqp/AA1980.DAT:*/SIDS8/gcgdata/geneseq/geneseqp/AA1981.DAT:* /SIDS8/gcgdata/geneseq/geneseqp/AA1982. /SIDS8/gcgdata/geneseq/geneseqp/AA1983. A_Geneseq_1101:*

/SIDS8/gcgdata/geneseq/geneseqp/AA1989.DAT: /SIDS8/gcgdata/geneseq/geneseqp/AA1990.DAT: /SIDS8/gcgdata/geneseq/geneseqp/AA1991.DAT /SIDS8/gcgdata/geneseg/genesegp/AA1985.DAT: /SIDS8/gcgdata/geneseg/genesegp/AA1986.DAT: /SIDS8/gcgdata/geneseg/genesegp/AA1987.DAT: /SIDS8/gcgdata/geneseq/geneseqp/AA1984.DAT: /SIDS8/gcgdata/geneseq/geneseqp/AA1988.DAT:

/SIDS8/gcgdata/geneseg/genesegp/AA1992.DAT:*

SUMMARIES

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

	Description	D. radiodurans del	Neurospora crassa	N. crassa delta228	B. subtilis delta2	S. pombe delta228-	GST signal peptide	GST signal peptide	Class II AP endonu	Puromycin acetyl t	Murine p53-induced	C glutamicum prote
	ID	A,	AAW01618	AAY44502	AAY44503	AAY44499	AAY44500	AAY44498	AAR42817	AAW48651	AAE00207	AAG90354
	DB	21	17	21	21	21	21	21	14	19	22	22
	Query Aatch Length DB ID	294	929	626	626	371	900	828	270	199	915	497
ю	Query Match	100.0	16.7	15.7	15.7	15.1	15.1	15.1	6.2	6.1	6.1	5.9
	Score	1527	255.5	239.5	239.5	230.5	230.5	230.5	94.5	93	92.5	90
	Result No.	1	7	m	4	ហ	9	7	80	6	10	11

Corynebe Sorangil Anycolal Anycolal Anycolal Anycolal Murine of Mycomor Murine of Mycromor Micromor Human por Human por Plateno,	
AAB79113 AAV58573 AAW52844 AAW52844 AAW52849 AAW31045 AAR989013 AAR989013 AAR989013 AAN16401 AAW3726 AAW3726 AAW3726 AAW3726 AAW372665 AAW23720 AAW23720 AAW23720 AAW23720 AAW23720 AAW23720 AAW23720 AAW23720 AAW23720 AAW23720 AAW23720 AAW23720 AAW23720 AAW23720 AAW23720 AAW23720 AAW23720 AAW23720 AAW23700 AAW23700 AAW23700 AAW23700 AAW23700 AAW270662 AAW270662 AAW19980	AAP80802 AAP80807 AAW22357 AAG32418 AAG32416 AAG32416 AAG32416
20000000000000000000000000000000000000	
497 1421 1421 1421 1421 1537 1537 1537 1539 1546 160 1160 1160 1160 1160 1160 1160 116	22288222
$\begin{array}{c} \mathbf{v} \cdot $	
888 888 888 888 888 888 888 888 888 88	
	4 4 4 4 4 4 4 4 4 4 4 4 4 4 4 4 4 4 4

AAY44505 standard; Protein; 294 AA. AAY44505

ALIGNMENTS

AAY44505;

(first entry) 27-MAR-2000

D. radiodurans delta228-UV damage endonuclease. XX Delta228-UVDE, ultraviolet damage endonuclease, GST signal peptide; glutathione-S-transferase signal peptide; uvel-4-gene product; UV irradiation, DNA damage; UV radiation damage; photoproduct; abasic site; aplatinum diaduct; mismatched nucleotide pairing; KW (

nucleotide alkylation; skin cancer. X

Dienococcus radiodurans. W09963828-A1

16-DEC-1999

99WO-US12910. 08-JUN-1999;

98US-0088521. 99US-0134752. 08-JUN-1998; 18-MAY-1999;

(UYEM-) UNIV EMORY.

Avery AM; Kaur B, Doetsch PW,

WPI; 2000-116417/10

A new truncated ultraviolet damage endonuclease for treatment of skin

```
Sequence
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            AAY44502;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  AAY44502
        δ
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 QQ
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         qq
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            0
                                                                                                                    D. radiodurans. Delta228-UVDE is a 288 amino acid deletion of the N-terminal of the uvel+ gene product. This is expressed in frame with a GST leader sequence to generate a fusion protein. This provides stable endonuclease fragments for cleaving a double-stranded DNA.molecule that has a distorted structure resulting from UV radiation damage, a photoproduct, an abasic site, mismatched nucleotide pairing, aplatinum diaduct, an intercalated molecule or alkylation of a nucleotide. Uvelp can be used in compositions for internal or topical application and as a therapeutic agent for skin cancers.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               121 VRESSVRAMSAHARVMDGLGLARTPWNLLLLHGGKGGRGAELAALIPDLPDPVRLRLGLE 180
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 NDERAYSPAELLPICEATGTPLVFDAHHHVVHDKLPDQEDPSVREWVLRARATWQPPEWQ 240
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            181 nderayspaellpiceatgtplvfdahhhvvhdklpdqedpsvrewvlraratwqppewq 240
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 DNA repair enzyme; cyclobutane-type dimer; (6-4)-bound product; exposure; UV light; ultraviolet; suntan cream; prevention; treatment;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  61 YRLSSSLFPMLDLAGDDTGAAVLTHLAPQLLEAGHAFTDAGVRLLMHPEQFIVLNSDRPE 120
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           1 OLGLVCLTVGPEVRFRTVTLSRYRALSPAEREAKLLDLYSSNIKTLRGAADYCAAHDIRL 60
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             241 VVHLSNGIEGPQDRRHSHLIADFPSAYADVPQIEVEAKGKEEAIAALRLMAPFK 294
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        DB 21; Length 294;
                                                                                                    The present sequence is delta228-UV damage endonuclease from
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              DNA repair enzyme and related DNA - specifically recognises
                                                                                                                                                                                                                                                                                                                                                                                                                                                        100.0%; Score 1527; DB 21; Length 100.0%; Pred. No. 3e-159; Live 0; Mismatches 0; Indels
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 Neurospora crassa DNA repair enzyme.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           AAW01618 standard; Protein; 656 AA
                                                Claim 16; Page 60; 133pp; English.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               95JP-0094137.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              95JP-0094137
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  (first entry)
                                                                                                                                                                                                                                                                                                                                                                                                                                                              Query Match 100.0
Best Local Similarity 100.0
Matches 294; Conservative
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   (SHIS ) SHISEIDO CO LTD.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               WPI; 1996-512663/51.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         Neurospora crassa.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          N-PSDB; AAT58286
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            29-MAR-1995;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               29-MAR-1995;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                20-AUG-1997
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        skin cancer.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              15-OCT-1996.
                                                                                                                                                                                                                                                                                                                                                                                    Sequence
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     AAW01618
g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  ô
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    QQ
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         d
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      δ
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        δý
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 ACCOMMENS ACCOMM
```

```
The present sequence shows a DNA repair enzyme which specifically recognises a cyclobutane-type dimer and (6-4)-bound product induced, respectively from TT and TC sequences in DNA, by exposure to UV light. The enzyme cleaves these sequences at the 5'-side. The DNA repair enzyme can be used in, e.g. suntan cream for prevention and treatment of skin cancer. Conventional endonuclease acts on either the cyclobutane-type dimer or the (6-4)-bound product in damaged DNA, while the enzyme of the invention can act on both of them.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         167 ----PDLPDPVRLRIGLENDERAYSPAELLPICEATGTPLVFDAHHHVV-----HDKL 215
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   39; Gaps
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       56 HDIRLYRLSSSLFPMLDLAGDDTGAAVLTHLAPQLL-EAGHAFTDAGVRLLMHPEQFIVL 114
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       292 ygirflrlssemfp---fashpvhgyklapfasevlaeagrvaaelghrltthpggftgl 348
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                115 NSDRPEVRESSVRAMSAHARVMDGLGLARTPWN-----LLLLHGGKGGRGAELAALI--- 166
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 349 gsprkevvesairdleyhdel---lsllklpeggnrdavmiih--mgggfgdkaatlerf 403
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      404 krnyarlsqscknrlvlenddvgwtvhdllpvceelnipmvldyhhhnicfdpahlregt 463
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  216 PDQEDPSVREWVLRARATWQPPE-WQVVHLSNGIEG----PQDRR-HSHLIADFPSAYADV 270
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     Delta228-UVDE; ultraviolet damage endonuclease; GST signal peptide; dutathione-S-transferase signal peptide; uvel+ gene product; UV irradiation; DNA damage; UV rediation damage; photoproduct; abasic site; aplatinum diaduct; mismatched nucleotide pairing;
cyclobutane-type dimer and (6-4)-bound product induced by exposure
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              DB 17; Length 656;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      103; Indels
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              16.7%; Score 255.5; DB 1 31.8%; Pred. No. 5.3e-19;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              38; Mismatches
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        N. crassa delta228-UV damage endonuclease.
                                                                                                      Disclosure; Page 10-13; 18pp; Japanese.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      AAY44502 standard; Protein; 626 AA.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         nucleotide alkylation; skin cancer.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   271 PQIEVEAKGKEEAIAALRLMAPFK 294
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            ::||| ||:|:
521 -dlmieakdkegav--felmrtfk 541
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          Kaur B, Avery AM;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       98US-0088521.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   99WO-US12910.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          99US-0134752
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       (first entry)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  Conservative
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          Query Match
Best Local Similarity
Matches 84; Conserv
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  (UYEM-) UNIV EMORY.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 Neurospora crassa.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  WO9963828-A1.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               08-JUN-1999;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   27-MAR-2000
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          8-MAY-1999;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       08-JUN-1998;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       16-DEC-1999.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   Doetsch PW,
```

```
WPI; 2000-116417/10
                                                                                                                                                                                                                                                                                                                                                                                   cancers.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    AAY44499
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   RESULT
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            qq
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              δ
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         qq
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            õ
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      qq
       12;
                                                                                                                                                                                    Delta228-UVDE is a 288 amino acid deletion of the N-terminal of the uvel+ gene product. This is expressed in frame with a GST leader sequence to generate a fusion protein. This provides stable endonuclease fragments for cleaving a double-stranded DNA molecule that has a distorted structure resulting from UV radiation damage, a photoproduct, an abasic site, mismatched nucleotide pairing, aplatinum diaduct, an intercalated molecule or alkylation of a nucleotide. Uvelp can be used in compositions for internal or topical application and as a therapeutic agent for skin
                                                                                                                                                                   The present sequence is delta228-UV damage endonuclease from N. crassa.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          142 ARTPWN-----LLLLHGGKGGRGAELAALI------PDLPDPVRLRLGLENDERAYSPA 189
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       : | | | | | | | : : : 358 lklpeqqnrdavmlih--mggqfgddxatlerfkrnyarlsqscknrlvlenddvgwtvh 415
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     :|||:|| |:| |:| |:| | |:| | |:| | |:| | |:| | |:| | |:| | |:| | |:| | |:| | |:| | |:| | |:| | |:| | |:| | |:| | |:| | |:| |:| |:| |:| |:| |:| |:| |:| |:| |:| |:| |:| |:| |:| |:| |:| |:| |:| |:| |:| |:| |:| |:| |:| |:| |:| |:| |:| |:| |:| |:| |:| |:| |:| |:| |:| |:| |:| |:| |:| |:| |:| |:| |:| |:| |:| |:| |:| |:| |:| |:| |:| |:| |:| |:| |:| |:| |:| |:| |:| |:| |:| |:| |:| |:| |:| |:| |:| |:| |:| |:| |:| |:| |:| |:| |:| |:| |:| |:| |:| |:| |:| |:| |:| |:| |:| |:| |:| |:| |:| |:| |:| |:| |:| |:| |:| |:| |:| |:| |:| |:| |:| |:| |:| |:| |:| |:| |:| |:| |:| |:| |:| |:| |:| |:| |:| |:| |:| |:| |:| |:| |:| |:| |:| |:| |:| |:| |:| |:| |:| |:| |:| |:| |:| |:| |:| |:| |:| |:| |:| |:| |:| |:| |:| |:| |:| |:| |:| |:| |:| |:| |:| |:| |:| |:| |:| |:| |:| |:| |:| |:| |:| |:| |:| |:| |:| |:| |:| |:| |:| |:| |:| |:| |:| |:| |:| |:| |:| |:| |:| |:| |:| |:| |:| |:| |:| |:| |:| |:| |:| |:| |:| |:| |:| |:| |:| |:| |:| |:| |:| |:| |:| |:| |:| |:| |:| |:| |:| |:| |:| |:| |:| |:| |:| |:| |:| |:| |:| |:| |:| |:| |:| |:| |:| |:| |:| |:| |:| |:| |:| |:| |:| |:| |:| |:| |:| |:| |:| |:| |:| |:| |:| |:| |:| |:| |:| |:| |:| |:| |:| |:| |:| |:| |:| |:| |:| |:| |:| |:| |:| |:| |:| |:| |:| |:| |:| |:| |:| |:| |:| |:| |:| |:| |:| |:| |:| |:| |:| |:| |:| |:| |:| |:| |:| |:| |:| |:| |:| |:| |:| |:| |:| |:| |:| |:| |:| |:| |:| |:| |:| |:| |:| |:| |:| |:| |:| |:| |:| |:| |:| |:| |:| |:| |:| |:| |:| |:| |:| |:| |:| |:| |:| |:| |:| |:| |:| |:| |:| |:| |:| |:| |:| |:| |:| |:| |:| |:| |:| |:| |:| |:| |:| |:| |:| |:| |:| |:| |:| |:| |:| |:| |:| |:| |:| |:| |:| |:| |:| |:| |:| |:| |:| |:| |:| |:| |:| |:| |:| |:| |:| |:| |:| |:| |:| |:| |:| |:| |:| |:| |:| |:| |:| |:| |:| |:| |:| |:| |:| |:| |:| |:| |:| |:| |:| |:| |:| |:| |:| |:| |:| |:| |:| |:| |:| |:| |:| |:| |:| |:| |:| |:| |:| |:| |:| |:| |:| |:| |:| |:| |:| |:| |:| |:| |:| |:| |:| |:| |:| |:| |:| |:| |:| |:| |:| |:| |:| |:| |:| |:| |:| |:| |:| |:| |:| |:| |:| |:| |:| |:| |:| |:| |:| |:| |:| |:| |:| |:| |:| |:| |:| |:| |:| |:| |:| |:| |:| |:| |:| |:| |:| |:| |:| |:| |:| |:| |:| |:| |:| |:| |:| |:| |:| 
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                LAP----QLLEAGHAFTDAGVRLLMHPEQFIVLNSDRPEVRESSVRAMSAHARVMDGLGL 141
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           301 lapfasevlaeagrvaaelghritthpgqftqlgsprkevvesairdleyhdel---lsl 357
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        190 ELLPICEATGTPLVFDAHHHVV-----HDKLPDQEDPSVREWVLRARATWQPPE-WQV 241
                                                 A new truncated ultraviolet damage endonuclease for treatment of skin
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    87; Indels 39; Gaps
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       Delta228-UVDE; ultraviolet damage endonuclease; GST signal peptide;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          242 VHLSNGIEG---PQDRR-HSHLIADFPSAYADVPQIEVEAKGKEEAIAALRLMAPFK 294
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             glutathione-S-transferase signal peptide; uvel+ gene product;
UV irradiation; DNA damage; UV radiation damage; photoproduct;
abasic site; aplatinum diaduct; mismatched nucleotide pairing;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       DB 21; Length 626;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          Score 239.5; DB 2 Pred. No. 2.8e-17;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               B. subtilis delta228-UV damage endonuclease.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         34; Mismatches
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     nucleotide alkylation; skin cancer.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       AAY44503 standard; Protein; 626 AA
                                                                                                                                Claim 16; Page 59; 133pp; English.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  Avery AM;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          15.7%;
32.5%;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           98US-0088521.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              99WO-US12910.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  (first entry)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   Local Similarity 32.5
nes 77; Conservative
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               Kaur B,
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                (UYEM-) UNIV EMORY.
            WPI; 2000-116417/10.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      Bacillus subtilis
                                                                                                                                                                                                                                                                                                                                                                                                                                         626 AA;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     W09963828-A1
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 08-JUN-1999;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    18-MAY-1999;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  27-MAR-2000
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                08-JUN-1998;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    16-DEC-1999
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    Doetsch PW,
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      AAY44503;
                                                                                                                                                                                                                                                                                                                                                                                                                                              Sequence
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 Query Match
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          98
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              Matches
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             RESULT
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            Ω
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           ŏ
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       셤
```

```
12;
                                                                                                                                       The present sequence is delta228-UV damage endonuclease from B. subtilis. Delta228-UVDE is a 288 amino acid deletion of the N-terminal of the uvel+ gene product. This is expressed in frame with a GST leader sequence to generate a fusion protein. This provides stable endonuclease fragments for cleaving a double-stranded DNA molecule that has a distorted structure resulting from UV radiation damage, a photoproduct, an abasic site, mismatched nucleotide pairing, aplatinum diaduct, an intercalated molecule or alkylation of a nucleotide. Uvelp can be used in compositions for internal or topical application and as a therapeutic agent for skin
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 142 ARTPWN-----LILLHGGKGGRGAELAALI------PDLPDPVRLRLGLENDERAYSPA 189
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        358 lklpeqqnrdavmiih--mggqfgdkaatlerfkrnyarlsgscknrlvlenddvgwtvh 415
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                190 ELLPICEATGTPLVFDAHHHVV-----HDKLPDGEDPSVREWVLRARATWQPPE-WQV 241
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     86 LAP----QLLEAGHAFTDAGVRLLMHPEQFIVLNSDRPEVRESSVRAMSAHARVMDGLGL 141
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      301 lapfasevlaeagrvaaelghritthpgqftqlgsprkevvesairdleyhdel---isi 357
A new truncated ultraviolet damage endonuclease for treatment of skin
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             39; Gaps
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                Delta228-UVDE; ultraviolet damage endonuclease; GST signal peptide;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 242 VHLSNGIEG----PQDRR-HSHLIADFPSAYADVPQIEVEAKGKEEAIAALRLMAPFK 294
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      glutathione-S-transferase signal peptide; uvel+ gene product; UV irradiation; DNA damage; UV radiation damage; photoproduct; abasic site; aplatinum diaduct; mismatched nucleotide pairing; nucleotide alkylation; skin cancer.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         15.7%; Score 239.5; DB 21; Length 626; 32.5%; Pred. No. 2.8e-17;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             Indels
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             87;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             34; Mismatches
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              pombe delta228-UV damage endonuclease.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     AAY44499 standard; Protein; 371 AA
                                                                                               Claim 16; Page 59; 133pp; English.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         Kaur B, Avery AM;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     98US-0088521.
99US-0134752.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                99WO-US12910
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          (first entry)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   Schizosaccharomyces pombe.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                Conservative
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                Query Match
Best Local Similarity
From 77; Conserve
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 (UYEM-) UNIV EMORY.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     626 AA;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                08-JUN-1999;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       18-MAY-1999;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              W09963828-A1
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           08-JUN-1998;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          27-MAR-2000
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         Doetsch PW,
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        16-DEC-1999
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     AAY44499;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           Sequence
```

```
AAY44498
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        RESULT
                     THE BENEVALUE OF STREET AND STREE
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             Οy
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            qq
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            ŏ
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            a
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                ολ
                                                                                                                                                                                           The present sequence is delta228-UV damage endonuclease. Delta228-UVDE is a 288 amino acid deletion of the N-terminal of the S. pombe uvel+ gene product. This is expressed in frame with a GST leader sequence to generate a fusion protein. This provides stable endonuclease fragments for cleaving a double-stranded DNA molecule that has a distorted structure resulting from UV radiation damage, a photoproduct, an abasic site, mismatched nucleotide pairing, aplatinum diaduct, an intercalated for internal or topical application and as a therapeutic agent for skin
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  10;
                                                                    A new truncated ultraviolet damage endonuclease for treatment of skin
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    59 RLYRLSSSLFPMLDLAGDDTGAAVLTHLAPQLLEAGHAFTDAGVRLLMHPEQFIVLNSDR 118
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     81 hfmrvssdlfp---fashakygytlefaqshleevgklankynhrltmhpgqytqiaspr 137
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   119 PEVRESSVRAMSAHARVMDGLGLARTPWN------LLLLHGGKGGRGAE-----LAALI 166
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     167 PDL.PDPVRLRLGL.ENDERAYSPAELL.PICEATGTPLVFDAHHHVV------HDKLPD 217
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     218 QEDPSVREWVLRARATWQPPEWQVVHLS----NGIEGPQDRRHSHLIADFPSAYADVPQ 272
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    1 QLGLVCLTVGPEVRFRTVTLSRYRALSPAEREA--KLLDLYSSNIKTLRGAADYCAAHDI 58
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  21 rigyacintilrsmkervfcsrtcrittigrdglesvkglgtgnvldliklvewnhnfgi 80
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             GST signal peptide; glutathione-S-transferase signal peptide; ultraviolet damage endonuclease; UVDE; delta228-UVDE; repair-deficient E. coli strain; UV irradiation. DNA damage; UV radiation damage; photoproduct; abasic site; aplatinum diaduct; mismatched nucleotide pairing; nucleotide alkylation;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           DB 21; Length 371;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                GST signal peptide and delta228 S. pombe UVDE fusion protein.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        42; Mismatches 140; Indels
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     15.1%; Score 230.5; DB 2 27.9%; Pred. No. 1.3e-16;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           Location/Qualifiers
                                                                                                                                              Claim 13; Page 53; 133pp; English.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            AAY44500 standard; Protein; 600 AA.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            fusion protein; skin cancer.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     88; Conservative
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 27-MAR-2000 (first entry)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         Schizosaccharomyces pombe
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     304 lmieakekegavfel 318
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        273 IEVEAKGKEEAIAAL 287
  WPI; 2000-116417/10.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                Similarity
                                                                                                                                                                                                                                                                                                                                                                                                                                                                         Sequence 371 AA;
                          N-PSDB; AAZ29859
                                                                                                    cancers -
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     Query Match
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      Synthetic
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 AAY44500;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                Best Local
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     Matches
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          AAY44500
δ
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    q
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      δ
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      δλ
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     qq
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      pp
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      ŏ
```

```
from S. pombe. Delta228-UVDE is a 288 amino acid deletion of the N-terminal of the S. pombe uvel+ gene product. This is expressed in frame with a GST leader sequence. Stable endonuclease fragments can be produced for cleaving a double-stranded DNA molecule that has a distorted structure resulting from UV radiation damage, a photoproduct, an abasic site, mismatched nucleotide pairing, aplatinum diaduct, an intercalated molecule or alkylation of a nucleotide. Uvelp can be used in compositions for internal or topical application and as a therapeutic agent for skin
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  new truncated ultraviolet damage endonuclease for treatment of skin
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               310 hfmrvssdlfp---fashakygytlefaqshleevgklankynhrltmhpgqytqiaspr 366
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      119 PEVRESSVRAMSAHARVMDGLGLARTPWN-----LLLLHGGKGGRGAE-----LAALI 166
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  250 rigyacintilrsmkervfcsrtcrittigrdglesvkqlgtqnvldliklvewnhnfgi 309
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              59 RLYRLSSSI,FPML,DLAGDUTGAAVLTHLAPQLLEAGHAFTDAGVRLLMHPEQFIVLNSDR 118
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           367 evvvdsairdlayhdei----lsrmklneqlnkdavliihlggtfegkketldrfrkny 421
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  167 PDLPDPVRLRLGLENDERAYSPAELLPICEATGTPLVFDAHHHVV------HDKLPD 217
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           218 QEDPSVREWVLRARATWQPPEWQVVHLS----NGIEGPQDRRHSHLIADFPSAYADVPQ 272
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                Indels 45; Gaps
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  1 QLGLVCLTVGPEVRFRTVTLSRYRALSPAEREA--KLLDLYSSNIKTLRGAADYCAAHDI 58
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        The present sequence is a fusion protein comprising the GST signal peptide and delta228 truncated ultraviolet damage endonuclease (UV
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        DB 21; Length 600;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   42; Mismatches 140;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            15.1%; Score 230.5; DB 2 27.9%; Pred. No. 2.6e-16;
/note= "Encoded by aaaa"
                                                    /note= "Encoded by gtt"
                                                                                                     /note= "Encoded by gtt"
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       Disclosure; Page 56; 133pp; English.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                        Avery AM.
                                                                                                                                                                                                                                                                        99WO-US12910
                                                                                                                                                                                                                                                                                                                            98US-0088521
                                                                                                                                                                                                                                                                                                                                                          99US-0134752
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   Conservative
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         273 IEVEAKGKEEAIAAL 287
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          533 Imieakekegavfel 547
                                                                                                                                                                                                                                                                                                                                                                                                                                                               Doetsch PW, Kaur B,
                                                                           Misc-difference 147
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    WPI; 2000-116417/10.
                                                                                                                                                                                                                                                                                                                                                                                                           (UYEM-) UNIV EMORY.
                    Misc-difference 85
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      Similarity
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           600 AA;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 N-PSDB; AAZ29860
                                                                                                                                                          WO9963828-A1
                                                                                                                                                                                                                                                               08-JUN-1999;
                                                                                                                                                                                                                                                                                                                            08-JUN-1998;
                                                                                                                                                                                                                                                                                                                                                       .8-MAY-1999;
                                                                                                                                                                                                                    16-DEC-1999
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              88;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              Sequence
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            Query Match
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        cancers
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         cancers.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         Local.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 Matches
```

218 QEDPSVREWVL,RARATWQPPEWQVVHL,S----NGIEGPQDRRHSHLIADFPSAYADVPQ 272

qq

qq ŏλ qq Οy

538 hfmrvssdlfp---fashakygytlefaqshleevgklankynhrltmhpgqytqiaspr 594

```
A new truncated ultraviolet damage endonuclease for treatment of skin
                                                              GST signal peptide; ultraviolet damage endonuclease; UVDE; Uvelp; UV irradiation; DNA damage; UV radiation damage; fusion protein; skin cancer; glutathione-S-transferase.
                                               GST signal peptide and S. pombe UVDE fusion protein.
                                                                                                                                            'label= GST_signal_peptide
                                                                                                                                                                           note= "Encoded by aaaa"
                                                                                                                                                                                            'note= "Encoded by gtt"
                                                                                                                                                                                                           /note= "Encoded by gtt"
                                                                                                                                                          'note= "S. pombe UVDE"
                                                                                                                           Location/Qualifiers
AAY44498 standard; Protein; 828 AA.
                                                                                                                                                                                                                                                                                                                                                                                        Claim 7; Page 51; 133pp; English.
                                                                                                                                                                                                                                                                                                                        Avery AM;
                                                                                                                                                                                                                                                            99WO-US12910
                                                                                                                                                                                                                                                                              98US-0088521
99US-0134752
                                  27-MAR-2000 (first entry)
                                                                                                   Schizosaccharomyces pombe
                                                                                                                                                    230..828
                                                                                                                                                                                                                                                                                                                       Doetsch PW, Kaur B,
                                                                                                                                                                                                                                                                                                                                       WPI; 2000-116417/10.
                                                                                                                                                                                                     Misc-difference 147
                                                                                                                                                                                                                                                                                                       UYEM-) UNIV EMORY.
                                                                                                                                                                                    Misc-difference 85
                                                                                                                                                                                                                                                                                                                                                N-PSDB; AAZ29858
                                                                                                                                                                      Misc-difference
                                                                                                                                                                                                                                                                                       18-MAY-1999;
                                                                                                                                                                                                                                                             08-JUN-1999;
                                                                                                                                                                                                                                                                                08-JUN-1998;
                                                                                                                                                                                                                             W09963828-A1
                                                                                                                                                                                                                                              16-DEC-1999
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             Sequence
                                                                                                            Synthetic
                                                                                                                                                                                                                                                                                                                                                                           cancers
                   AAY44498
                                                                                                                                    Peptide
                                                                                                                                                    Protein
```

Apurinic/apyrimidinic; recombinant; heat stable.

Thermus thermophilus

WO9320191-A. 14-OCT-1993. 93WO-US03035.

30-MAR-1993; 31-MAR-1992; 16-APR-1992;

92US-0869306 92US-0860861

(CHAD) CHAKAKI DYEING KK.

(ABBO) ABBOTT LAB

Backman KC, McLean PA;

WPI; 1993-336899/42.

N-PSDB; AAQ49396

AAR42817 standard; Protein; 270 AA.

AAR42817

273 IEVEAKGKEEAIAAL 287 : :||| ||:|: | 761 Imieakekegavfel 775 27-APR-1994 (first entry)

AAR42817;

Class II AP endonuclease.

```
stranded DNA molecule that has a distorted structure resulting from UV radiation damage, a photoproduct, an abasic site, mismatched nucleotide pairing, aplatinum diaduct, an intercalated molecule or alkylation of a nucleotide. Uvelp can be used in compositions for internal or topical application and as a therapeutic agent for skin cancers.
The present sequence is a fusion protein comprising the GST signal peptide and full length ultraviolet damage endonuclease (UVDE) from Schizosaccharomyces pombe. UVDE genes are isolated by introducing a foreign cDNA library into a repair-deficient E. coli strain and selecting for complemented cells by UV irradiation of the transformants. This provides stable endonuclease fragments for cleaving a double-
```

```
1 QLGLVCLTVGPEVRFRTVTLSRYRALSPAEREA--KLLDLYSSNIKTLRGAADYCAAHDI 58
15.1%; Score 230.5; DB 21; Length 828; 27.9%; Pred. No. 4.2e-16;
                                            Indels
                                            140;
                                            42; Mismatches
  Ouery Match
Best Local Similarity 27.99
Matches 88; Conservative
```

10;

59 RLYRLSSSLFPMLDLAGDDTGAAVLTHLAPQLLEAGHAFTDAGVRLLMHPEQFTVLNSDR 118

> g ò

ò

```
17;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        Gaps
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             22 RYRALSPAEREAKLLDLYSSNIKTLRGAADYCAA--HDIRLYRLSS--SLF--PMLDLAG 75
                                                                                                                                                                                                                                                                                                                                                     (apurinic/apyrimidinic) endonuclease which retains activity when subjected to elevated temperatures for the time necessary to effect denaturation of double-stranded nucleic acids. It is useful
                                                                                                                                                                                                                                                                                                                                                                                      in ligase chain reactions for nucleic acid sequence amplification and may be used to "correct" the ends of probes that have been modified to prevent blunt end ligation in absence of the target, but only when the probe is hybridised to the target.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       67; Indels 89;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                               6.2%; Score 94.5; DB 14; Length 270; 27.5%; Pred. No. 0.068;
                                                                                                                                                                                                                                                                                                                                           sequence is that of Thermus thermophilus class II AP
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           Mismatches
                                                                                                                                                                                                                                                                                                                        Claim 16; Fig 1B; 42pp; English.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           29;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            Conservative
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                Best Local Similarity
                                                                                                                                                                                                                                                                                                                                                                                                                                                     270 AA;
                                                                                                                                                                                                                                                                                                                                                                                                                                                      Sequence
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        Query Match
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             Matches
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     δò
```

New recombinant DNA encoding heat stable AP endonuclease - and derived vectors, host cells and proteins, useful for nucleic acid amplification by ligase chain reaction

```
The present sequence represents a puromycin acetyl transferase protein which was used as a selection marker. The invention claims for a new pMCLDHAP triciastronic vector (AAV18096) for the expression of an antibody-cytokine fusion protein, hmAb425-THF alpha. hmAb425-THF alpha comprises of the TWF alpha fused to the C-terminus of the heavy for humanized monoclonal antibody 425. The hmAb425 has specificity by the IL-2 sequence. The TWF alpha sequence can be substituted by the IL-2 sequence. The vector also contains a strong promoter/ derived internal ribosomal entry site (IRES) sequences. The vector can derived internal ribosomal entry site (IRES) sequences. The vector can be expressed in mammalian host cells for the production of heteromeric fusion proteins. This expression system is claimed to produce the
                                         76 DDTGAAVLTHLAPQLLEAGHAFTDAGVR-LLMHPEQFIVLNSDRPE-VRESSVRAMSAHA 133
                                                                                                                      93 dlekaail-----gyeyvvvhp-----gsgrpervkegalkalrl-- 127
                                                                                                     134 RVMDGLGLARTPWNLLLLH----GGK--GGRGAELAALIPDLPDPVRLRLGLENDERAYS 187
                                                                                                                                                               188 PAELLPICEATGTPLVFDAHHHVVHDKLPDQEDPSVREWVLRARATWQPPEWQVVHLSNG 247
                                                                                                                                                                                  Oligo:cistronic expression vector - useful for production of, e.g. MAD425/TNF-^a or MAD425/IL-2 antibody fusion protein
                                                                                                                                                                                                                                                                                                                                                                                                                                                  Antibody-cytokine fusion protein; tricistronic vector; TNF alpha; IL-2; IRES; internal ribosome entry site.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             Hauser H, Mielke C;
                                                                                                                                                                                                                                                                                                                              AAW48651 standard; Protein; 199 AA.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      Disclosure; Fig 15; 89pp; English.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           P. Burge C, Dunker R,
Von Hoegen I, Welge T;
                                                                                                                                                                                                                                                                                                                                                                                                                     Puromycin acetyl transferase.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     97WO-EP04765.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               96EP-0114820.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   96EP-0115635
                                                                                                                                                                                                                                                                                                                                                                                         04-AUG-1998 (first entry)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          (MERE ) MERCK PATENT GMBH.
                                                                                                                                                                                                                                               209 vgglgsrvdhhahll 223
                                                                                                                                                                                                                         248 IEGPQDR--RHSHLI 260
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          Burge C,
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                WPI; 1998-207400/18.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                199 AA;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    N-PSDB; AAV18096
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          WO9811241-A1.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   02-SEP-1997;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                30-SEP-1996;
16-SEP-1996;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       19-MAR-1998
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          Bruemmer W,
                                                                                                                                                                                                                                                                                                                                                           AAW48651;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        Rieke E,
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  Seguence
                                                                                                                                                                                                                                                                                                                AAW48651
                                             ò
                                                                        qq
                                                                                                      δŏ
                                                                                                                                 qq
                                                                                                                                                               ò
                                                                                                                                                                                          qq
                                                                                                                                                                                                                         δ
                                                                                                                                                                                                                                                      Ω
```

```
13;
                                                                                                                                                                                                                                                                                                                                                                                                                Murine; immunosuppressive; neuroprotective; nootropic; cerebroprotective; cytostatic; apoptosis; p53-induced protein; death domain; PIDD; therapy; cell cycle regulation; autoimmune disease; stroke; dysplasia; cancer;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       The present sequence is a murine p53-induced protein with a death domain
                                                                                                       88 PQLLEAGHAFTDAGVRLLMHPEQFIVLNSDRPEVRESSVRAMSAHARVMDGLGLARTP-- 145
                                                                                                                               79 pesveagaviaeig------prmaelsgsrlaaggg-megilaphrpke 120
                                  Gaps
                                                                                                                                                          146 --WNLLLL-----HGGKGGRGAELAALIPDLPDPVRLRLGLENDERAYSPAEL----- 191
                                                                                                                                                                             43 IKTLRGA-ADYCAA------HDIRLYRLSSSLFPMLDL-----AGDDTGAAVLTHLA 87
                                                                    Novel p53 induced protein with a death domain that promotes apoptosis, useful for treating cancer, dysplasia, autoimmune disease, neurodegenerative disorders and stroke
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             /note= "N-terminal 7 tandem leucine-rich repeat"
6.1%; Score 93; DB 19; Length 199; 26.0%; Pred. No. 0.063; ative 18; Mismatches 65; Indels
                                                                                                                                                                                                                                                                                                                                                                                        Murine p53-induced protein with a death domain (PIDD).
                                                                                                                                                                                                               192 LPICEATGTPLVFDAHHHVVHDKLPDQEDPSVREWVLRARATW 234
                                                                                                                                                                                                                                         168 lpfyerlgftvtadv-----ecp-----kdratw 191
                                                                                                                                                                                                                                                                                                                                                                                                                                                          neurodegenerative disorder; Li-Fraumeni syndrome.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        /label= Antigenic_fragment_#1
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     824.838
/label= Antigenic_fragment_#2
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            /label= Death_domain
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 Location/Qualifiers
                                                                                                                                                                                                                                                                                                          AAE00207 standard; Protein; 915 AA
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               Claim 13; Fig 2a; 85pp; English.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  (UYHE-) UNIV HEALTH NETWORK.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   /label= LRR
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        07-SEP-1999; 99US-0152359.
13-JUN-2000; 2000US-0211086.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                07-SEP-2000; 2000WO-CA01021.
              Best Local Similarity 26.0%
Matches 58; Conservative
                                                                                                                                                                                                                                                                                                                                                               13-JUN-2001 (first entry)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            .877
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         .291
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          Benchimol S, Lin Y;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   WPI; 2001-244557/25.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  N-PSDB; AAD03265
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            WOZ00118037-A2
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      15-MAR-2001.
 Query Match
                                                                                                                                                                                                                                                                                                                                     AAE00207;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     Mus sp.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           Region
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     Region
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            Domain
                                                                                                                                                                                                                                                                                RESULT 10
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      Region
                                                                                                                                                                                                                                                                                             AAE00207
                                                    δ
                                                                            qq
                                                                                                                              QQ
                                                                                                       ŏ
                                                                                                                                                            δ
                                                                                                                                                                                    qq
                                                                                                                                                                                                                                                                                                                                    δλ
                                                                                                                                                                                                                                       q
```

```
The present invention provides a number of nucleotide and protein sequences from the Coryneform bacterium Corynebacterium glutamicum. These are useful for identifying the mutation point of a gene derived from a mutant of coryneform bacterium, measuring expression amount and analysing the expression profile or expression pattern of a gene derived
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 Novel polynucleotides derived from Coryneform bacteria, for identifying mutation point of a gene, measuring expression of a gene, analysing expression profile or pattern of a gene and identifying homologous gene
of prob) that induces apoptosis. p53 can hold cells in check, quiescent and prevent cells from becoming cancerous. PIDD comprises leucine-rich responsive to p53. PIDD is useful for treating or preventing diseases associated with cell cycle regulation, cell growth and deficient or insufficient apoptosis, autoimmune disease, neurodegenerative disorder,
                                                                                                                                                                                                                                                                                                                                            128 AMSAH--ARVMDGLGLARTPWNL----LLLHGG--KGGRGAELAALIPDLPDPVRLRLG 178
                                                                                                                                                                                                                                                                                                                                                                 80 AAVLTHLAPQILEA-----GHAFTDAGVRLLM------HPEQFIVLNSDRPEVRESSVR 127
                                                                                                                                                                                                 6.1%; Score 92.5; DB 22; Length 915;
26.7%; Pred. No. 0.72;
Live 29; Mismatches 62; Indels 27; Gaps
                                                                                                                                                                                                                                                                                                        14 aaaaedaatstleavdagpgapflpagnqlnldlrpggchrlqyl-csqqppqllqvefl 72
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             Ando S, Hayashi M, Ochiai K, Yokoi H;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      Coryneform bacterium; amino acid synthesis; vitamin; saccharide;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             Claim 17; SEQ ID NO: 4108; 246pp + Sequence Listing; English.
                                                                                                                                                                                                                                                                                                                                                                                                                  179 LENDERAYSPAELLPIC--EATGTPLVFDAHHHVVHDKLPD 217
                                                                                                                                                                                                                                                                                                                                                                                                                                           stroke, cancer, dysplasia and Li-Fraumeni syndrome.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      C glutamicum protein fragment SEQ ID NO: 4108.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 Mizoguchi H, Ando S, Hayas
Senoh A, Ikeda M, Ozaki A;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   AAG90354 standard; Protein; 497 AA
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  (KYOW ) KYOWA HAKKO KOGYO KK.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            L8-DEC-2000; 2000EP-0127688.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            99JP-0377484.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              03-AUG-2000; 2000JP-0280988
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 2000JP-0159162
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     Corynebacterium glutamicum.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          26-SEP-2001 (first entry)
                                                                                                                                                                                                                                Local Similarity 26.7% les 43; Conservative
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     organic acid synthesis.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           WPI; 2001-376931/40.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              N-PSDB; AAH65573
                                                                                                                                                           915 AA;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          Nakagawa S,
Tateishi N,
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  16-DEC-1999;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   07-APR-2000;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         EP1108790-A2
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           20-JUN-2001.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         AAG90354;
                                                                                                                                                             Sequence
                                                                                                                                                                                                                  Query Match
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     RESULT 11
                                                                                                                                                                                                                                                     Matches
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          AAG90354
```

g

ò

õ

88888888888

g ŏ

```
fine chemical production; organic acid; proteinogenic amino acid; nonproteinogenic amino acid; purine base; pyrimidine base; nucleoside; nucleotide; lipid; saturated fatty acid; unsaturated fatty acid; dial; carboydrate; aromatic compound; vitamin; cofactor; polyketide; enzyme; diagnosis; Corynebacterium diphtheriae; genetic engineering; Brevibacterium; environmental condition.
from coryneform bacterium, and identifying a homologue of a gene derived from coryneform bacterium. Coryneform bacteria are useful for producing amino acids, nucleic acids, vitamins, saccharides and organic acids, particularly L-lysine. The present sequence is a protein described in the exemplification of the invention.

Note: The sequence data for this patent did not form part of the printed specification, but was obtained in electronic format directly from the
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             275 admsvlstr1-----irscm---rnt-----gqtcyistriiapssr------ 308
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              146 WNLLLLHGGKGGRGAELAALIPDLPDPVRLRLGLENDERAYSPAELLPICEATGTPLVFD 205
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 232 gvdkvaftgstpvgkkiaaacgellrpvtlelg-----gkssaillp-----d 274
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    206 AHHHVVHDKLPDQEDPSVREWVLRARATWQPPEWQVVHLSNGIEGPQDRRHSHLIADFPS 265
                                                                                                                                                                                                                                                                                                                                                                                                   124 ryfatlapwid--gedirpfpagsaesivdkdpigvcaliapwnfpinlvviklapali- 180
                                                                                                                                                                                                                                                                                                                                                                                                                                       93 AGHAFTDAGVRLIMHPE-----QFIVLNSDRPEVRESSVRAMSAHARVMDGLGLARTP 145
                                                                                                                                                                                                                                                                                                                                                                                                                                                                        181 -----agctviikpasptplsirfiieaieaagvpagvvnlltgsgrfgd--alvrhp 231
                                                                                                                                                                                                                                                                                                                             68 ikswsaltgaertgylikiatelesrsealaltntrengspisetrgaasnaag----if 123
                                                                                                                                                                                                                                                              Gaps
                                                                                                                                                                                                                                                                                                                                                                    ---VLTHLAPQLLE 92
                                                                                                                                                                                                                                                                                             -----SNIKTLRGAADYCAAHDIRLY 61
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   Corynebacterium glutamicum; homeostasis; adaptation; HA protein;
                                                                                                                                                                                                                     5.9%; Score 90; DB 22; Length 497;
23.5%; Pred. No. 0.54;
ive 31; Mismatches 90; Indels 114;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                Corynebacterium glutamicum HA protein sequence SEQ ID NO:182.
                                                                                                                                                                                                                                                                                                                                                                          62 RLSSSLFPMLDLAGDD-----TGAA-----
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              AAB79113 standard; Protein; 497 AA.
                                                                                                                                                                                                                                                                                                      20 LSRYRALSPAEREAKLLDLYS-----
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        99DE-1032125.
99DE-1032126.
99DE-1032127.
99DE-1032128.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   23-JUN-2000; 2000WO-IB00911.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           99DE-1031636.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        99US-0141031.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             Corynebacterium glutamicum.
                                                                                                                                                                                                                                              Best Local Similarity 23.5%;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           30-APR-2001 (first entry)
                                                                                                                                                                                                                                                                       72; Conservative
                                                                                                                                          European Patent Office.
                                                                                                                                                                                 497 AA;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             309 -yaevvg 314
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          266 AYADVPQ 272
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   WO200100842-A2
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            09-JUL-1999;
09-JUL-1999;
09-JUL-1999;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              09-JUL-1999;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      04-JAN-2001.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           25-JUN-1999;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 08-JUL-1999;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          AAB79113;
                                                                                                                                                                                   Sednence
                                                                                                                                                                                                                                       Query Match
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     RESULT 12
                                                                                                                                                                                                                                                                           Matches
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        AAB79113
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              qq
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              οy
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 qq
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  Ŏλ
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    Op
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         δò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             qq
                                                                                                                                                                                                                                                                                                                                                                                                                                                          ΟŊ
                       8*888888888
```

```
AAF71138 to AAF71357 encode the Corynebacterium glutamicum homeostasis

C and adaptation (HA) proteins given in AAB79023 to AAB7924.

C glutamicum HA genes (I) can be used in vectors for expression in host

C cells and production of films chemicals, such as, an organic acid,

proteinogenic or nonproteinogenic amino acid (preferred), purine or

C faity acid, diol, carbohydrate, aromatic compound, vitamin, cofactor,

C fatty acid, diol, carbohydrate, aromatic compound, vitamin, cofactor,

C fatty acid, diol, carbohydrate, aromatic compound, vitamin, cofactor,

C polyketide or enzyme. The amino acids produced can be lysine, glutamine,

C glutamate, aladine, aspartate, glycine, serine, threonine, methionine,

C cysteine, valine, leucine, isoleucine, arginine, proline, histidine,

C tyrosine, phenylalanine, or tryptophan. The fine chemical production can

cused for diagnosing the presence of (I) or HA proteins encoded by then are

C diphtheriae. (I) can be used to map the C. glutamicum genome or can be

C used as markers for genetically engineered Corynebacterium or

C bemeostasis in C. glutamicum or help the microorganism to adapt to
                                                                                                                                                                                                                                                                                                                                                                                                                        New isolated Corynebacterium glutamicum nucleic acid for production or modulation of production of fine chemicals such as amino acids, nucleosides, nucleotides, lipids, fatty acids, carbohydrates, vitamins or enzymes -
                                                                                                                                                                                                                                                                                                                                             Pompejus M, Kroeger B, Schroeder H, Zelder O, Haberhauer G;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                Claim 20; Page 390-391; 712pp; English.
      99DE-1032129
                                   99DE-1032920
                                                   99DE-1032922
                                                                    99DE-1032924
                                                                                  99DE-1032928
                                                                                               99DE-1032930
                                                                                                                 99DE-1032933.
                                                                                                                               99DE-1032935
                                                                                                                                                           99DE-1033002.
                                                                                                                                                                               .033003
                                                                                                                                                                                         99DE-1033005.
                                                                                                                                                                                                                                        99DE-1041379.
                                                                                                                                            99DE-1032973
                                                                                                                                                                                                                         99DE-1041378
                                                                                                                                                                                                                                                                                      99DE-1042088
                                                                                                                                                                                                                                                                        99DE-1041391
                                                                                                                                                                              99DE-1
                                                                                                                                                                                                                                                                                                                                                                           WPI; 2001-061974/07.
                                                                                                                                                                                                                                                                                                                (BADI ) BASF AG.
                                                                                                                                                                                                                                                                                                                                                                                            N-PSDB; AAF71228
                                                                                                                                         14-JUL-1999;
14-JUL-1999;
14-JUL-1999;
                                                                                                 6661
                                                  14-JUL-1999;
                                 14-JUL-1999;
                                                                                                                                                                                                                                                                                 03-SEP-1999;
                                                                                                                                                                                                                                   31-AUG-1999;
                                                                                                                                                                                                                      31-AUG-1999
                                                                             14-JUL-1
14-JUL-1
14-JUL-1
14-JUL-1
```

14; Ouery Match Best Local Similarity 23.5%; Pred. No. 0.54; Matches 72; Conservative 31; Mismatches 90; Indels 114; Gaps 68 lkswsaltgaertgyllkiateiesrsealaltntrengspisetrgaasnaag----if 123 | ::| | || |:| | 124 ryfatlapwld--gedirpfpagsaesivdkdpigvcaliapwnfpinlvviklapall- 180 93 AGHAFTDAGVRLLMHPE-----OFIVLNSDRPEVRESSVRAMSAHARVMDGLGLARTP 145 181 -----agctviikpasptplsirfiieaieaagvpagvvnlltgsgrfgd--alvrhp 231 146 WNLLLLHGGKGGRGAELAALIPDLPDPVRLRLGLENDERAYSPAELLPICEATGTPLVFD 205 20 LSRYRALSPAEREAKLLDLYS------SNIKTLRGAADYCAAHDIRLY 61 62 RLSSSLFPMLDLAGDD-----TGAA----------------------VLTHLAPQLLE 92 Query Match ò ò CD δ q

497 AA;

Sequence

Oy 206 ANIHIVATIONE, PROCESSAY KIRAGONGEN INTELLINGENESS 26: 206 ANIHIVATIONE, PROCESSAY KIRAGONGEN INTELLINGENESS 26: 206 ANIHIVATIONE, PROCESSAY KIRAGONGEN INTELLINGENESS 26: 206 ANADAYO 275 D) 275 AUMEN'SET1

```
used for the recombinant production of epothilones in a heterologous host that is more amenable to fermentation.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  Amycolatopsis mediterranei strain wt3136 5.7 kb KpnI fragment protein.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       178 GLENDERAYSPAELLPICEATGTP-----LVFDAHHHVVHDKLPDQE-DPSV-- 223
                                                                                                                                                                                                                                                                                                                                                                                                                                 340 glgrd--vatplligsvktnlghpeyasgitgllkvvlsIghgqipahlhaqalnprisw 397
                                                                                                                                                                                                                                                                                                                                                                                                                                                                   224 ---REWVLRARATWQPPEWQVVH----LSNGIEGPQDRRHSHLI-ADFPSAYADVPQIEV 275
                                                                                                                                                                                                                                       113 VLNSDRPEVRESSVRAMSAHARVMDGLGLARTPWNLLL--LHGGK--GGR------ 158
                                                                                                                                                                                                                                                                                                              223 wisktralardgrckafsaea---dgfgrgegcavvvlkrlsgaradgdrilavirgsai 279
                                                                                                                                                                                                                                                                                                                                             280 nhdgassgltvpngssqeivlkraladagcaassvgyveahgtgttlgdpieiqalnavy 339
                                                                                                                                                                                                                   55 AHDIRLYRLSSSLFPMLDLAGDDTGAAVLTHLAPQLLEAGHAFT--DAGVRLLMHPEQFI 112
                                                                                                                                                                                      129 vflgigp----seyeaalpqatasaeidahgglgtmpsvgagrisyalglrgpcv 179
                                                                                            Ouery Match

5.9%; Score 90; DB 21; Length 1421;
Best Local Similarity 21.5%; Pred. No. 2.6;
Matches 78; Conservative 37; Mismatches 122; Indels 126; Gaps
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           Amycolatopsis mediterranei rifamycin synthesis gene cluster – used
to produce rifamycin and rifamycin analogues
                                                                                                                                                       5 VCLTVGPEVRFRTVTLSRYRALSPAEREAKLLDLY-----SSNIKTLRGAADYCA 54
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                Amycolatopsis mediterranei; rifamycin; synthesis; gene cluster; polyketide synthase; actinomycete; ansamycin.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              Example 6; Page 43-53; 205pp; English.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              AAW52844 standard; Protein; 1891 AA.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       Engel N, Schupp T, Toupet C;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            96EP-0810551.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               97WO-EP04495
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      Amycolatopsis mediterranei
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              24-JUL-1998 (first entry)
                                                                                                                                                                                                                                                                                                                                                     159 ---GAELAALIPD------
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     WPI; 1998-169172/15.
N-PSDB; AAV21186.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           (NOVS ) NOVARTIS AG.
                                                            1421 AA;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    18-AUG-1997;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       WO9807868-A1
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               20-AUG-1996;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         452 pae 454
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    AAW52844;
                                                              Sequence
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         RESULT 14
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         AAW52844
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 q
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               οy
                                                                                                                                                                                                                                                                                                                                                                                                                                                     g
                                                                                                                                                                                                                                                                                                                                                                                                                         õ
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    ò
```

QQ δλ g

ò g δ

δ

S X X S

```
The present sequence represents Amycolatopsis mediterranei strain wt3136

5.7 kb KpnI fragment protein, from the present invention. The present
invention describes a Amycolatopsis mediterranei rifamycin synthesis.

convention describes a Amycolatopsis mediterranei rifamycin synthesis.

convention describes a Amycolatopsis mediterranei rifamycin synthesis.

including the adjacent DNA regions to the right and left which, by

reason of their function in connection with rifamycin blosynthesis,

reason of their function in connection with rifamycin blosynthesis.

constituents of this rifamycin gene cluster, and functional

qualify as constituents of this rifamycin gene cluster. DNA fragment can be used

for producing rifamycin synthesis gene cluster DNA fragment can be used for constructing mutant

consideration in the DNA can be used for constructing mutant

continomycetes strains from which the natural rifamycin or ansamycin

considerate can be used for assembling a library of polyketides

synthases, which can be used for assembling a library of polyketides

consideration probe of the invention can be used for identifying DNA

configuration probe of the invention can be ansamycins.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  10;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                1023 rlagiapaeqealildvvrtqv----al 1046
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       199 GTPLVFDAHHHVVHDKLPDQEDPSVREWVLRARATWQPPE--WQVVHLSNGIEG----PQ 252
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            142 ARTPWNLLLLHGGKGGRGAELAALIPDLPDPVRLRLGLEN---DERAYSPAELLPICEAT 198
                                                                                                                                                                                                                                                                                                                                                                                                                                                         Query Match 5.9%; Score 90; DB 19; Length 1891;
Best Local Similarity 22.3%; Pred. No. 4.1;
Matches 54; Conservative 21; Mismatches 61; Indels 106; Gaps
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  82 VLTHLAPQLLEAGHAFTDAGVRLLMHPEQFTVLNSDRPEVRESSVRAMSAHARVMDGLGL 141
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        A. mediterranei rifamycin synthesis gene cluster fragment protein E.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     22 RYRALSPAEREAKLIDLYSSNIKTLRGAADYCAAHDTRLYRLSSSLFPMLDLAGDDTGAA 81
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 Amycolatopsis mediterranei; rifamycin; synthesis; gene cluster;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      polyketide synthase; actinomycete; ansamycin.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                AAW52849 standard; Protein; 3413 AA.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      97WO-EP04495
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         96EP-0810551
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  Amycolatopsis mediterranei.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 (first entry)
                                                                                                                                                                                                                                                                                                                                                                                                                 1891 AA;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               1127 gepi----
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               20-AUG-1996;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           WO9807868-A1.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  26-FEB-1998.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       1166 dr 1167
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     24-JUL-1998
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    253 DR 254
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                AAW52849;
                                                                                                                                                                                                                                                                                                                                                                                                                     Seguence
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      RESULT 15
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        AAW52849
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               δλ
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            qq
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 QΥ
          οy
```

(NOVS) NOVARTIS AG.

```
Engel N, Schupp T, Toupet C;

XX

WPI: 1998-169172/15.

R N-PSDB; AAV21187.

XX

Amycolatopsis mediterranei rifamycin synthesis gene cluster - used
XX

Amycolatopsis mediterranei rifamycin synthesis gene cluster - used
XX

Caim 6; Page 170-187; 205pp; English.

C Synthesis gene cluster ORF E protein from the present invention. The
DNA fragment comprises a DNA region involved directly or indirectly
CC shift fragment comprises a DNA region involved directly or indirectly
CC in the gene cluster responsible for rifamycin synthesis, including
CC in the adjacent DNA regions to the right and left which, by reason of
CC in the gene cluster of this rifamycin gene cluster, and functional
CC cas constituents of this rifamycin gene cluster. The Amycolatopsis
CC for producing rifamycin, synthesis gene cluster. DNA fragment can be used
CC for producing rifamycin synthesis gene cluster DNA fragment can be used
CC actinomycetes strains from which the natural rifamycin or ansamycin or
CC actinomycetes strains from which the natural rifamycin or ansamycin or
CC actinomycetes strains from which the natural rifamycin or ansamycin on
CC actinomycetes strains from which the natural rifamycin or ansamycin or
CC actinomycetes strains from which the natural rifamycin or ansamycin or
CC A hybridisation probe of the invention can be used for assembling a library of polyketides
CC A hybridisation probe of the invention can be used for identifying DNA

XX

Sequence 3413 AA;
```

----- 1620 1662 p-----atlvfdyptpvalarylrdelgdtvattpvataaaada-1700 61; Indels 106; Gaps ------alvgmacrlpggvtdpeglwrlvr--dgleglspfpe 1739 82 VLTHLAPQLLEAGHAFTDAGVRLLMHPEQFIVLNSDRPEVRESSVRAMSAHARVMDGLGL 141 142 ARTPWNLLLLHGGKGGRGAELAALIPDLPDPVRLRLGLEN---DERAYSPAELLPICEAT 198 199 GTPLVFDAHHHVVHDKLPDQEDPSVREWVLRARATWQPPE--WQVVHLSNGIEG----PQ 252 22 RYRALSPAEREAKLLDLYSSNIKTLRGAADYCAAHDIRLYRLSSSLFPMLDLAGDDTGAA 81 Query Match 5.9%; Score 90; DB 19; Length 3413; Best Local Similarity 22.3%; Pred. No. 9.9; Matches 54; Conservative 21; Mismatches 61; Indels 10 1701 gepi-----1740 dr 1741 253 DR 254 Óγ δλ Dp δλ Q δλ

Search completed: January 15, 2002, 13:50:18 Job time: 307 sec

1

```
January 15, 2002, 13:50:57; Search time 57.72 Seconds (without alignments) 233.532 Million cell updates/sec
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        Appl
Appl
Appl
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     Appli
Appli
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  Appli
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          Appli
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   3, Appli
2, Appli
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                Appli
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             Appli
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                Sequence 4, Appli
Sequence 4, Appli
Sequence 37, Appl
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            Sequence 23, App]
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        Sequence 17,
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                Sequence 29,
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           Sequence 16,
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       Sequence 30,
                                                                                                                                                                                                                                                                                                                                      1 MLRLLKRNIQISKRIVFTIL.......KRLTARKRRSRKEEVEEDEK 599
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 Sequence 11,
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                Sequence 1,
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              Sequence 39,
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     Sequence 40,
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        Seguence 1,
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             Sequence 2,
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           Sequence 3,
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  Sequence 2,
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        Sequence 5,
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         Description
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     Sequence
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      Sequence
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                Sequence
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     Sequence
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   Sequence
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                Sequence
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            Sequence
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      Sequence
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             Sequence
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      Sequence
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        /cgn2_6/ptodata/2/iaa/5A_COMB.pep:*
/cgn2_6/ptodata/2/iaa/5B_COMB.pep:*
/cgn2_6/ptodata/2/iaa/6A_COMB.pep:*
/cgn2_6/ptodata/2/iaa/6B_COMB.pep:*
/cgn2_6/ptodata/2/iaa/PcTUS_COMB.pep:*
/cgn2_6/ptodata/2/iaa/PcTUS_COMB.pep:*
GenCore version 4.5
Copyright (c) 1993 - 2000 Compugen Ltd.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     PCT - US93 - 0.7261 - 1.1
PCT - US93 - 0.7261 - 1.1
PCT - US93 - 0.7261 - 1.6
US - 0.8 - 9.23 - 9.92.4
US - 0.8 - 9.23 - 9.92.4
US - 0.8 - 7.25 - 4.59.8 - 3.7
US - 0.8 - 7.25 - 4.59.8 - 3.8
US - 0.8 - 7.25 - 4.59.8 - 3.8
US - 0.8 - 7.25 - 4.59.8 - 3.9
US - 0.8 - 7.25 - 4.59.8 - 3.9
US - 0.8 - 7.25 - 4.59.8 - 3.9
US - 0.8 - 7.25 - 4.59.8 - 3.9
US - 0.8 - 7.25 - 4.59.8 - 3.9
US - 0.8 - 7.25 - 8.36.2 - 7.0
US - 0.8 - 7.25 - 8.36.2 - 7.0
US - 0.8 - 7.25 - 8.36.2 - 7.0
US - 0.8 - 7.25 - 8.36.2 - 7.0
US - 0.8 - 7.25 - 7.0
US - 0.8 - 7.0
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             US-08-725-459B-29
US-08-725-459B-30
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              US-08-725-459B-23
US-08-725-459B-25
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      US-08-725-459B-26
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          US-08-081-929-2
US-08-725-459B-3
US-08-725-459B-2
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         Total number of hits satisfying chosen parameters:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 212252 seqs, 22503292 residues
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       SUMMARIES
                                                                                                                                                                                                                                                                                 US-09-724-296-2_COPY_230_828
3112
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           Maximum Match 100%
Listing first 45 summaries
                                                                                                                  OM protein - protein search, using sw model
                                                                                                                                                                                                                                                                                                                                                                                                    BLOSUM62
Gapop 10.0 , Gapext 0.5
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      Issued_Patents_AA:*
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 seq length: 0
seq length: 200000000
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   Post-processing: Minimum Match 0%
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         Match Length DB
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            Query
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              140.5
134.5
134.5
133.5
133.5
133.5
130.5
127.5
127.5
127.5
127.5
127.5
127.5
127.5
127.5
127.5
127.5
127.5
127.5
127.5
127.5
127.5
127.5
127.5
127.5
127.5
127.5
127.5
127.5
127.5
127.5
127.5
127.5
127.5
127.5
127.5
127.5
127.5
127.5
127.5
127.5
127.5
127.5
127.5
127.5
127.5
127.5
127.5
127.5
127.5
127.5
127.5
127.5
127.5
127.5
127.5
127.5
127.5
127.5
127.5
127.5
127.5
127.5
127.5
127.5
127.5
127.5
127.5
127.5
127.5
127.5
127.5
127.5
127.5
127.5
127.5
127.5
127.5
127.5
127.5
127.5
127.5
127.5
127.5
127.5
127.5
127.5
127.5
127.5
127.5
127.5
127.5
127.5
127.5
127.5
127.5
127.5
127.5
127.5
127.5
127.5
127.5
127.5
127.5
127.5
127.5
127.5
127.5
127.5
127.5
127.5
127.5
127.5
127.5
127.5
127.5
127.5
127.5
127.5
127.5
127.5
127.5
127.5
127.5
127.5
127.5
127.5
127.5
127.5
127.5
127.5
127.5
127.5
127.5
127.5
127.5
127.5
127.5
127.5
127.5
127.5
127.5
127.5
127.5
127.5
127.5
127.5
127.5
127.5
127.5
127.5
127.5
127.5
127.5
127.5
127.5
127.5
127.5
127.5
127.5
127.5
127.5
127.5
127.5
127.5
127.5
127.5
127.5
127.5
127.5
127.5
127.5
127.5
127.5
127.5
127.5
127.5
127.5
127.5
127.5
127.5
127.5
127.5
127.5
127.5
127.5
127.5
127.5
127.5
127.5
127.5
127.5
127.5
127.5
127.5
127.5
127.5
127.5
127.5
127.5
127.5
127.5
127.5
127.5
127.5
127.5
127.5
127.5
127.5
127.5
127.5
127.5
127.5
127.5
127.5
127.5
127.5
127.5
127.5
127.5
127.5
127.5
127.5
127.5
127.5
127.5
127.5
127.5
127.5
127.5
127.5
127.5
127.5
127.5
127.5
127.5
127.5
127.5
127.5
127.5
127.5
127.5
127.5
127.5
127.5
127.5
127.5
127.5
127.5
127.5
127.5
127.5
127.5
127.5
127.5
127.5
127.5
127.5
127.5
127.5
127.5
127.5
127.5
127.5
127.5
127.5
127.5
127.5
127.5
127.5
127.5
127.5
127.5
127.5
127.5
127.5
127.5
127.5
127.5
127.5
127.5
127.5
127.5
127.5
127.5
127.5
127.5
127.5
127.5
127.5
127.5
127.5
127.5
127.5
127.5
127.5
127.5
127.5
127.5
127.5
127.5
127.5
127.5
127.5
127.5
127.5
127.5
127.5
127.5
127.5
127.5
127.5
127.5
127.5
127.5
127.5
127.5
127.5
127.5
127.5
127.5
127.5
127.5
127.5
127.5
127.5
127.5
127.5
127.5
127.5
127.5
127.5
127.5
127.5
127.5
127.5
127.5
127.5
127.5
127.5
127.5
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              125.5
125.5
125.5
125.5
                                                                                                                                                                                                                                                                                                                     Perfect score:
                                                                                                                                                                                                                                                                                                                                                                                                       Scoring table:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    Minimum DB
Maximum DB
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         Database :
                                                                                                                                                                                                                                                                                                                                                  Sequence:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          Searched:
                                                                                                                                                                                Run on:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               .
9
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  Result
```

```
15;
                                                                                                                           5, Appli
5, Appli
2, Appli
30, Appl
1, Appli
                 79, Appl
1, Appli
                                                                                      sequence 24, Appl
Patent No. 5183745
                                                                                                          Appl
                                     Appl
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              Gaps
               Sequence 79, App
Sequence 1, App
Sequence 22, App
Sequence 13, App
Sequence 13, App
                                                                                                         Sequence 27,
                                                                                                                    Sequence 18,
                                                                                                                            Sequence
Sequence
Sequence
Sequence
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              83;
                                                                             Sequence
                                                                                                                                                                      Sequence
                                                                   Sequence
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       4.5%; Score 140.5; DB 3; Length 733; 20.1%; _Pred. No. 0.00083;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              Indels
                                                                                                                                                                                                                                                                                   APPLICANT: CONAWAY, RONALD C.
APPLICANT: CONAWAY, JOAN W.
TITLE OF INVENTION: ELONGIN A AND C FUNCTIONAL DOMAINS
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        /note= "amino acids 1-480 and 521-773 of Elongin A"
                                                                                                                                                                                                                                                                                                                                                                                                                            CURRENT APPLICATION DATA:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             Mismatches 167;
                                   US-08-725-4599-22
US-08-205-21-13
US-08-961-810-123
US-08-352-9020-123
US-08-255-836C-66
US-08-725-4598-24
5183745-3
                                                                                                         US-08-725-459B-27
US-08-339-152A-18
US-08-007-999B-5
US-08-689-276A-5
US-08-123-702-2
                                                                                                                                                             US-08-339-152A-30
        US-08-725-459B-1
US-08-725-459B-79
                             PCT-US93-03077-1
                                                                                                                                                                        US-08-104-165-1
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            11146/07501
                                                                                                                                                                                                                                                                                                                                                                                                                                                             APPLICATION NUMBER: US/08/725,459B
                                                                                                                                                                                                                                                                                                                                                 STREET: 717 N. HARWOOD, SUITE 3400
                                                                                                                                                                                                                                                      ; Sequence 23, Application US/08725459B
; Patent No. 6084068
; GENERAL INFORMATION:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                75;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      REGISTRATION NUMBER: 31,966
REGISTRATION NUMBER: 31,966
REFERENCE/DOCKET NUMBER: 111
TELECOMMUNICATION INFORMATION:
TELEPHONE: 214-981-3300
                                                                                                                                                                                                                                                                                                                                                                                                                         IBM PC compatible
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              not relevant
                                                                                                                                                                                                                                                                                                                                        SIDLEY & AUSTIN
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            TELEFAX: 214-981-3400
INFORMATION FOR SEQ ID NO: 23:
                                                                                                                                                                                                                                                                                                                                                                                                    COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            ATTORNEY/AGENT INFORMATION:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          733 amino acids
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         not relevant
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 Conservative
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           TELEPHONE: 214-50.
TELEPHONE: 214-981-3400
meterAX: 214-981-3400:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               SEQUENCE CHARACTERISTICS:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         TOPOLOGY: not relevar
MOLECULE TYPE: protein
773
773
1093
770
770
770
770
770
748
800
748
694
695
                                                                                                                                                                                                                                                                                                                               CORRESPONDENCE ADDRESS:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     OTHER INFORMATION:
                                                                                                                                                                                                                                                                                                                      NUMBER OF SEQUENCES:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          OTHER INFORMATION:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       NAME/KEY: Protein
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    amino acid
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              Query Match
Best Local Similarity
Matches 82; Conserv
                                                                                                                                                                                                                                                                                                                                                                                            75201-6507
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    CLASSIFICATION:
DALLAS
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               STRANDEDNESS:
                                                                                                                                                                                                                                                                                                                                                                                  SD
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               US-08-725-459B-23
                                                                                                                                                                                                                                                US-08-725-459B-23
                                                                                                                                                                                                                                                                                                                                                                                                                          COMPUTER:
                                                                                                                                                                                                                                                                                                                                           ADDRESSEE:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   LOCATION
                                                                                                                                                                                                                                                                                                                                                                                  COUNTRY:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            LENGTH:
STATE:
                                                                                                                                                                                                                                      RESULT
 228
331
331
332
332
333
34
44
44
45
45
45
```

```
133 DSSTHLRRSSRSKKPVN--YNSSSESEEQISKATKKVKQKEEEEYVEEVDEKSLKNES 190
                                                                                                 73 NSFSAETPVDLKKENETELANISGPHKKSTSTSTRKRARSSKKKATDSVSDKIDESVASY 132
                                                                                                                                              191 SSDEFEPV-----VPEQLETPISKRRS--RSSAKNLEKESTMNLDDHAPR-----EMF 237
                                                                                                                                                                               347 GRGTGDPLPRAKDKVPNNLKAQEGKVRTNSDRKSPGSLPKVEEMDMDDEFEQPTMSFESY 406
                                                                                                                                                                                                                     238 DCLDKPIPWRGRLGYACLNTILRSMKERVFCSRTCRITTIQRDGLESVKQLGT-QNVLDL 296
                                                                                                                                                                                                                                              -----SHAKYGYTLEF 332
                                                                                                                                                                                                                                                                                                                    333 AQSHLEEVGKLANKYNHRLTMHPGQYTQIASPREVVVDSAIRDLAYHDEILSRMKLNEQL 392
                                                                                                                                                                                                                                                                                                                                                                                                    508 MTLHQQCIRVLKNNIDSIFEVGGVPYSVL----EPVLERCTPDQLYRIEECNHVLIEE-- 561
                                                                                                                                                                                                                                                                                                                                                                                                                                        393 NKDAVLIIHLGGTF-EGKKETLDRFRKNYQRLSDSVKARLVLENDDV 438
                                                                                                                                                                                                                                                                                                                                                                                                                                                             ELONGIN A AND C FUNCTIONAL DOMAINS
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          NAME/KEY: Protein
LOCATION: 1..753
OTHER INFORMATION: /note= "amino acids 1-545 and
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              SOFTWARE: Patentin Release #1.0, Version #1.30 CURRENT APPLICATION DATA:
                                                                                                                                                                                                                                                                                             297 IKLVEWNHNFGIHFM------RVSSDLFPFA--
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     APPLICAT: CONTACT ELONGIA:
APPLICANT: CONTACT ELONGIA:
NUMBER OF SEQUENCES: 79
CORRESPONDENCE ADDRESS:
ADDRESSEE: SIDLEY & AUSTIN
"TEFT: 71" HARMOOD, SUITE 3400
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               11146/07501
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                UMBER: US/08/725,459B
04-OCT-1996
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    Sequence 25, Application US/08725459B Patent No. 6084068 GENERAL INFORMATION:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 PC-DOS/MS-DOS
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              IBM PC compatible
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       NAME: HANSEN, EUGENIA S.
REGISTRATION NUMBER: 31,966
REFERENCE/DOCKET NUMBER: 111
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            TELECOMMUNICATION INFORMATION:
TELEPHONE: 214-981-3300
TELEFAX: 214-981-3400
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            INFORMATION FOR SEO ID NO: 25:
SEQUENCE CHARACTERISTICS:
LENGTH: 753 amino acids
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          SS: not relevant not relevant
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 FILING DATE: 04-OCT-1996
CLASSIFICATION: 530
ATTORNEY/AGENT INFORMATION:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      MOLECULE TYPE: protein
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    APPLICATION NUMBER:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            COMPUTER: IBM PC OPERATING SYSTEM:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       TYPE: amino acid
STRANDEDNESS: not
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   COUNTRY: US
ZIP: 75201-6507
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               US-08-725-459B-25
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           TOPOLOGY:
                                                                        δ
                                                                                                        g
                                                                                                                                              δŏ
                                                                                                                                                                                  qq
                                                                                                                                                                                                                       ò
                                                                                                                                                                                                                                                       Q
                                                                                                                                                                                                                                                                                             δ
                                                                                                                                                                                                                                                                                                                             g
                                                                                                                                                                                                                                                                                                                                                                    ò
                                                                                                                                                                                                                                                                                                                                                                                                    g
                                                                                                                                                                                                                                                                                                                                                                                                                                          δ
```

```
73 NSFSAETPVDLKKENETELANISGPHKKSTSTSTRKRARSSKKKATDSVSDKIDESVASY 132
                                                                                                                                                                                        133 DSSTHLRRSSRSKKPVN--YNSSSESESEQISKATKKVKQKEEEEYVEEVDEKSLKNES 190
                                                                                                                                                                                                                                                                                297 EKD----REGNSLKKKLSPALDVASDNHFKKPKHKDSEKIKSDKNKQSVDSVD-----S 346
                                                                                                                                                                                                                                                                                                                             191 SSDEFEPV-----VPEQLETPISKRRS--RSSAKNLEKESTMNLDDHAPR-----EMF 237
                                                                                                                                                                                                                                                                                                                                                                    347 GRGTGDPL?RAKDKVPNNLKAQEGKVRTNSDRKSPGSLPKVEEMDMDDEFEQPTMSFESY 406
                                                                                                                                                                                                                                                                                                                                                                                                             238 DCLDKPIPWRGRLGYACLN"ILRSMKERVFCSRTCRITTIQRDGLESVKQLGT-QNVLDL 296
                                                                                                                              95; Gaps
                                                                                                                                                                                                                                                                                                                                                                                                                                                    407 LSYDQP------RKKKKKVVKTSG---TALGEKGLKKKDSKSTSKNLNSA 447
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            297 IKLVEWNHNFGIHFM------RVSSDLFPF-----ASHAKYG--YTLEFAQS-- 335
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         377 AYHDEILSRMKLNEQLNKDAVLIIHLGGTF-EGKKETLDRFRKNYQRLSDSVKARLVLEN 435
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          ----HLEEVGKLANKYNHRLTMHPGQ---YTQIA-----SPREVVVDSAIRDL 376
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   508 PKRKAFSSPQEEEEAGFTGRRMNSKMQVYSGSKCAYLPFEVGGVPYSVLBPVLERCTPDQ 567
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   568 LYRIEBCNHVLIEE---TDQLWKVHCHRDFKEERPEEYESWREMYLRLQDAREQRLKLLT 624
                                                                             4.3%; Score 134.5; DB 3; Length 753; 20.3%; Pred. No. 0.0029; IVE 73; Mismatches 169; Indels 95
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            APPLICANT: CONAWAY, RONALD C.
APPLICANT: CONAWAY, JOAN W.
TITLE OF INVENTION: ELONGIN A AND C FUNCTIONAL DOMAINS
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       SOFTWARE: Patentin Release #1.0, Version #1.30 CURRENT APPLICATION DATA:
OTHER INFORMATION: 566-773 of Elongin A"
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            11146/07501
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       CORRESPONDENCE ADDRESS:
CORRESPONDENCE SIDLEY & AUSTIN
STREET: 717 N. HARWOOD, SUITE 3400
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    JMBER: US/08/725,459B
04-OCT-1996
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  Sequence 26, Application US/08725459B Patent No. 6084068
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        PC-DOS/MS-DOS
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    REGISTRATION NUMBER: 31,966
REPERDICE/DOCKET UNMBER: 111
TELECOMMUNICATION INFORMATION:
TELEPHONE: 214-981-3300
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  IBM PC compatible
                                                                                                    20.3%;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             ZIP: 75201-6507
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
                                                                                                                        86; Conservative
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            ATTORNEY/AGENT INFORMATION:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             HANSEN, EUGENIA S
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             NUMBER OF SEQUENCES: 79
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  TELEFAX: 214-981-3400
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  REGISTRATION NUMBER:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             APPLICATION NUMBER:
FILING DATE: 04-OCT
                                                                             Query Match
Best Local Similarity
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  COMPUTER: IBM PC
OPERATING SYSTEM:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         GENERAL INFORMATION:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          STREET: /...
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        CLASSIFICATION:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     Sn
    ; US-08-725-459B-25
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           436 DDV 438
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 US-08-725-459B-26
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               625 NNI 627
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   COUNTRY:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 NAME:
                                                                                                                        Matches
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           RESULT
                                                                                                                                                              δŽ
                                                                                                                                                                                                     q
                                                                                                                                                                                                                                              δλ
                                                                                                                                                                                                                                                                                     qq
                                                                                                                                                                                                                                                                                                                             ò
                                                                                                                                                                                                                                                                                                                                                                                                           QΥ
                                                                                                                                                                                                                                                                                                                                                                                                                                                    qq
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 QQ
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            Qγ
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            δō
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   a
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  qq
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           à
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           δò
```

```
Sequence 16, Application PC/TUS9307261
GENERAL INFORMATION:
TITLE OF INVENTION: PEEMP3 MALARIA ANTIGEN, ANALOGS, ANTIBODIES AND USES THEREO
                                                                                                                                                                                                                                                                                                                                                                                      26;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            261 SMKERVFCSRTCRITTLQRDGLESVKQLGTQNVLDLIKLVEWNHNFGIHFMRVSSDLFPF 320
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     321 ASHAKYGYTLEFAQSHLEEVGKLANKYNHRLTMHPGQYTQIASPREVVVDSAIRDLAYHD 380
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                ---EKLNAQLESAINELKE--RRASRRPMMVKMQRGMKDE-----VDEWLK--KYDD 210
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           381 EILSRMKLNEQLNKDAVLIIHLGGTFEGKKET------LDRFFRKNYQRLSDSVK 428
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         211 EQAEKNGTKDEEIKDK-----GDGYEEIVETKFYGMRENALGELDEYEERYE-----K 258
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       429 ARLVLENDDVSWSVQDLLPLCQELNIPLVLDWHHHNIVPGTLREGSLDLMPLIPTIRETW 488
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    259 KRYYLK-EDGEGDLKD--------VEFKLEETGYGFREKFPTTRILV 296
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  489 TRKGITQKQHYSESADPTAISG------MKRRAHSDRVF------DFPPCDPTMDL 532
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             297 KRKRNKEQKKLKEDKEKKLIAAEEPDDEKKIKLKDSDDKVVVPVNKNKSSFPD----- 349
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           533 MIEAKEKEQAVFELCRRYELQN--PPCP-----LEIMGPEYDQTRDGYYPPGAEKRLTA- 584
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        350 KFRAPDKKRTWF-----YRLSELFPLVPRKDNELAVCGDSMDSKVNG------KKLKST 397
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   201 BQLETPISKRRSRSSAKNLEKESTMNLDDHAPREMFDCLDKPIPWRGRLGYACLNTILR 260
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              83 LKKENETELANISGPHKKSTSTRKRARSSKKKATDSVSDKIDESVASYDSSTHLRRSS 142
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          143 RS-KKPVNYNSSSESESEEQ1-SKATKKVKQKEEEETVEEVDEKSLKNESSSDEFEPVVP 200
                                                                                                                                                                                                                                                                                                                                                                                      Indels 191; Gaps
                                                                                                                                                                                                                                                                                                                                                                                                                                                                               7 LKQKTE-----KNEKARNALKEKKLKEQKKNDAQKAKDLTKKE--SQDSS----SE 51
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       153 SLR------RDKLREIEQL--------
                                                                                                                                                                                                                                                                                                                                         Ouery Match 4.3%; Score 133.5; DB 5; Best Local Similarity 19.4%; Pred. No. 0.011; Matches 108; Conservative 84; Mismatches 175;
                                                DX0288K
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   John H. C. Blasdale
NAME: Blasdale, John H. C. REGISTRATION NUMBER: 31,895
                                           REFERENCE/DOCKET NUMBER: DX TELECOMMUNICATION INFORMATION:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         STREET: One Giralda Farms
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           585 ----RKRRSRKEEVEEDE 598
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      398 FNPFKRRRNKLKERKMQE 415
                                                                                                                                                                                      1588 amino acids
                                                                                           : 201-822-7398
201-822-7039
                                                                                                                                        INFORMATION FOR SEQ ID NO: SEQUENCE CHARACTERISTICS:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  CORRESPONDENCE ADDRESS:
                                                                                                                                                                                                                                                       MOLECULE TYPE: protein
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          NUMBER OF SEQUENCES:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           New Jersey
                                                                                                                                                                                                            amino acid
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      07940-1000
                                                                                                                                                                                                                                    linear
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      Madison
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 USA
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            RESULT 5
PCT-US93-07261-16
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          ADDRESSEE:
                                                                                           TELEPHONE:
                                                                                                                                                                                                                                                                                 PCT-US93-07261-11
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              COUNTRY:
                                                                                                                                                                                                                                 TOPOLOGY:
                                                                                                                     TELEFAX:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      CITY: 1
STATE:
                                                                                                                                                                                      LENGTH:
                                                                                                                                                                                                            TYPE:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           qa
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        qq
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            qq
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            δλ
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     qa
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                δy
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              q
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            δ
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  Ob
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       δλ
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                qq
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                ŏλ
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    Οy
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               qq
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        ŏ
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   Op
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          TITLE OF INVENTION: PÉEMP3 MALARIA ANTIGEN, ANALOGS, ANTIBODIES AND USES THEREOF
                                                                                                                                                                                                                                                                                                                                                                                                  16;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 297 IKLVEWNHNFGIHFM------RVSSDLFPF-----ASHAKYG--YTLEFAQS-- 335
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              ------HLEEVGKLANKYNHRLTMHPGQYTQIASPREVVVDSAIRDLAYHDEILSR 385
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     386 MKLNEQLN-----KDAVLIIHLGGTF-EGKKETLDRFRKNYQRLSDSVKARLVLENDDV 438
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  568 LYRIEECNHVLIEETDQLWKVHCHRDFKEERPEEYESWREMYLRLQDAREQRLRLLTNNI 627
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      297 EKD---REGNSLKKKLSPALDVASDNHFKKPKHKDSEKIKSDKNKQSVDSVD-----S 346
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              407 LSYDQP------RKKKKKVVKTSG---TALGEKGLKKKDSKSTSKNLNSA 447
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       508 PKRKAFSSPQEEEEAGFTGRRMNSKMQVYSGSKCAYLPKMMTLHQQCIRVLKNNIDSIDQ 567
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      133 DSSTHLRRSSRSKKPVN--YNSSSESEEQISKATKKVKQKEEEEYVEEVDEKSLKNES 190
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               191 SSDEFEPV-----VPEQLETPISKRRS--RSSAKNLEKESTMNLDDHAPR-----EMF 237
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        238 DCLDKPIPWRGRLGYACLNTILRSMKERVFCSRTCRITTIQRDGLESVKQLGT-QNVLDL 296
                                                                                                                                                                                                                                                                                                                                                                                                                                             73 NSFSAETPVDLKKENETELANISGPHKKSTSTSTRKRARSSKKKATDSVSDKIDESVASY 132
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      243 SSHKEKRPVDARGDEKSSVMGREKSHKASSKEESRRLL-----SEDSAKEKLPSSVVKK 296
                                                                                                                                                                                                                                                                                                                                                                                                  Gaps
                                                                                                                                                                                                                                                                                                                                                     DB 3; Length 753;
                                                                                                                                                                                                                                                                                                                                                Query Match

4.3%; Score 134.5; DB 3; Length

Best Local Similarity 19.8%; Pred. No. 0.0029;

Matches 83; Conservative 75; Mismatches 173; Indels
                                                                                                                                                                                                                                        /note= "amino acids 1-565 and 586-773 of Elongin A"
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        APPLICATION NUMBER: PCT/US93/07261
FILING DATE: 19930805
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              MEDIUM TYPE: Floppy disk
comptyre: Apple macintosh
OPERATING SYSTEM: Macintosh 6.0.5
SOFTWARE: Microsoft Word 5.1a
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 07/927,531
FILING DATE: 07-AUG-1992
ATTORNEY AGENT INFORMATION:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 Sequence 11, Application PC/TUS9307261 GENERAL INFORMATION:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    John H. C. Blasdale
                                                                                                       not relevant
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            STREET: One Giralda Farms
            INFORMATION FOR SEQ ID NO: 26:
                                                         : 753 amino acids
amino acid
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       CURRENT APPLICATION DATA:
                                      SEQUENCE CHARACTERISTICS:
                                                                                                                               not relevant
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         NUMBER OF SEQUENCES: 2
CORRESPONDENCE ADDRESS:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           COMPUTER READABLE FORM:
                                                                                                                               TOPOLOGY: not relevan
                                                                                                                                                                                                                                        OTHER INFORMATION:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          New Jersey
                                                                                                                                                                                                NAME/KEY: Protein
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 07940-1000
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       Madison
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                USA
                                                                                                    STRANDEDNESS:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           RESULT 4
PCT-US93-07261-11
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         ADDRESSEE:
                                                                                                                                                                                                                                                                                         US-08-725-459B-26
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             COUNTRY:
                                                                LENGTH:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          STATE:
```

g ŏ a ò g õ g

ò

g ò

```
26;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               Query Match
4.3%; Score 133.5; DB 5; Length 1663;
Best Local Similarity 19.4%; Pred. No. 0.012;
Matches 108; Conservative 84; Mismatches 175; Indels 191; Gaps
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 83 LKKENETELANISGPHKKSTSTSTRKRARSSKKKATDSVSDKIDESVASYDSSTHLRRSS 142
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      143 RS-KKPVNYNSSSESESEEQI-SKATKKVKQKEEEEYVEEVDEKSLKNESSSDEFEPVVP 200
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               52 KSLKEKVNGEALKEKENKETLKKKELENQKEKEEKNKIKDNNDEALKNKGNDKDDKKIVP 111
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          201 EQLETPISKRRSRSSAKNLEKESTMNLDDHAPREMFDCLDKPIPWRGRLGYACLNTILR 260
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                112 KKPESVEKDLKEMELKEKEFIKQ---HLKDYEERK-----EKRRNW------ILR 152
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            261 SMKERVFCSRTCRITTIQRDGLESVKQLGTQNVLDLIKLVEWNHNFGIHFMRVSSDLFPF 320
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     153 SLR------RDKLREIEQL------165
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                321 ASHAKYGYTLEFAQSHLEEVGKLANKYNHRLTMHPGQYTQIASPREVVVDSAIRDLAYHD 380
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              381 EILSRMKLNEQLNKDAVLIIHLGGTFEGKKET------LDRFRKNYQRLSDSVK 428
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         211 EQAEKNGTKDEEIKDK-----GDGYEEIVETKFYGMRENALGELDEYEERYE-----K 258
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               259 KRYYLK-EDGEGDLKD------VEEKLEETGYGFREKFPTTRILV 296
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      429 ARLVLENDDVSWSVQDLLPLCQELNIPLVLDWHHHNIVPGTLREGSLDLMPLIPTIRETW 488
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          489 TRKGITQKQHYSESADPTAISG------MKRRAHSDRVF------DFPPCDPTMDL 532
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   297 KRKRNKEQKKLKEDKEKKLIAAEEPDDEKKIKLKDSDKVVVPVNKNKSSFPD----- 349
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          7 LKOKTE-----SE05S----SE51
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            533 MIEAKEKEQAVFELCRRYELQN--PPCP----LEIMGPEYDQTRDGYYPPGAEKRLTA- 584
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          OPERATING SYSTEM: Macintosh 6.0.5
SOFTWARE: Microsoft Word 5.1a
CURRENT APPLICATION DATA:
                                                                                                                                 APPLICATION NUMBER: PCT/US93/07261
                                                                                                                                                                                               APPLICATION NUMBER: US 07/927,531
FILING DATE: 07-AUG-1992
ATTORNEY_AGENT INFORMATION:
                                                                                                                                                                                                                                                            NAME: Blasdale, John H. C.
REGISTRATION NUMBER: 31,895
REFERNCE/DOCKET NUMBER: DX0288K
TREECOMMUNICATION INFORMATION:
TELEPHONE: 201-822-7398
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           ORGANISM: Plasmodium falciparum
                                                                                                                                                                                                                                                                                                                                                                                         INFORMATION FOR SEQ ID NO: 16: SEQUENCE CHARACTERISTICS:
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: Apple Macintosh
                                                                                                                                                                                                                                                                                                                                                                                                                                       1663 amino acids
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              398 FNPFKRRRNKLKERKMQE 415
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   585 ----RKRRSRKEEVEEDE 598
                                                                                                                                                      19930805
                                                                                                                                                                                                                                                                                                                                                                          201-822-7039
                                                                                                                                                   FILING DATE: 19930805
PRIOR APPLICATION DATA:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  MOLECULE TYPE: peptide ORIGINAL SOURCE:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    STRAIN: Malayan Camp
                                                                                                                                                                                                                                                                                                                                                                                                                                                         amino acid
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     linear
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      PCT-US93-07261-16
                                                                                                                                                                                                                                                                                                                                                                                                                                                                              TOPOLOGY:
                                                                                                                                                                                                                                                                                                                                                                                                                                       LENGIH:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      δλ
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        δλ
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                qq
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              à
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       qq
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    ð
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         q
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          δ
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               ά
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     g
```

```
TITLE OF INVENTION: STRESS-INDUCED PROTEINS, GENES CODING
TITLE OF INVENTION: THEREFOR, TRANSFORMED CELLS OF ORGANISMS, METHODS AND
TITLE OF INVENTION: APPLICATIONS
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           ٠.
ش
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      51 KSLLPMSSKTTLSMLPQVN1GANSFSAETPVDLKKENETELANISGPHKKSTSTST--- 106
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    107 ----RKRARSSKKKATDSVSDKIDESVASYDSSTHLRRSSRSKKPVNYNSSSESEEEQ 161
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            77 EAETKKEESKDSSSSSDSSSDEEEEEKEETKKEESKESSSDS--SSSSSDSEKE 134
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   56; Indels 20; Gaps
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               Sequence 4, Application US/08923992A
Patent No. 6280738
GENERAL INFORMATION:
APPLICANT: Tai Joseph Y.
APPLICANT: Blake, Milan S.
TITLE OF INVENTION: Streptococcal Beta Antigens
NUMBER OF SEQUENCES: 34
CORRESPONDENCE ADDRESS:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  Query Match 4.2%; Score 132; DB 1; Length 414; Best Local Similarity 26.3%; Pred. No. 0.0019; Matches 40; Conservative 36; Mismatches 56; Indels 2
                                                                                                                                                                                                                                                                                                                                                                                                                                          PatentIn Release #1.0, Version #1.25
                                                                                                                                                                                                                              E: Weiser & Associates
230 S. Fifteenth Street, Suite 500
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       162 ISKATKKVKQKEEEEYVEEVDEKSLKNESSSD 193
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            135 ESNDKKRKSEDAEEEEDEESSNKKOKNEETEE 166
                                                                                                                                                                                                                                                                                                                                                                                                                                                                               US/07/667,276A
                                                                                                                                                                                                                                                                                                                                                                                                                  OPERATING SYSTEM: PC-DOS/MS-DOS
US-07-667-276A-4; Sequence 4, Application US/07667276A; Patent No. 5470971
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          REFERENCE/DOCKET NUMBER: 377
TELECOMMUNICATION INFORMATION:
TELEPHONE: 215-875-8383
                                                                                                                                                                                                                                                                                                                                                                                                IBM PC compatible
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       NAME: Weiser, Gerard J. REGISTRATION NUMBER: 19,763
                                                             GENERAL INFORMATION:
APPLICANT: Kondo, Keiji
APPLICANT: Inouye, Masayori
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        ATTORNEY/AGENT INFORMATION:
                                                                                                                                                                                                                                                                                                                                                                             Floppy disk
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   11-MAR-1991
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           : 414 amino acids
amino acid
                                                                                                                                                                                                                                                                                                                                                                                                                                                     CURRENT APPLICATION DATA:
APPLICATION NUMBER: US
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     SEQUENCE CHARACTERISTICS:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              TELEFAX: 215-875-8394 INFORMATION FOR SEQ ID NO:
                                                                                                                                                                                                           CORRESPONDENCE ADDRESS:
                                                                                                                                                                                                                                                                                                                                                       COMPUTER READABLE FORM:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  MOLECULE TYPE: protein
                                                                                                                                                                                                                                                                        Philadelphia
                                                                                                                                                                                         NUMBER OF SEQUENCES:
                                                                                                                          TITLE OF INVENTION:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           linear
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    CLASSIFICATION:
                                                                                                                                                                                                                                                                                                                  USA
                                                                                                                                                                                                                                                                                                                                                                          MEDIUM TYPE:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     FILING DATE:
                                                                                                                                                                                                                                                                                                                                   19102
                                                                                                                                                                                                                                                                                            STATE: PA
                                                                                                                                                                                                                              ADDRESSEE:
                                                                                                                                                                                                                                                                                                                                                                                              COMPUTER:
                                                                                                                                                                                                                                                                                                                                                                                                                                          SOFTWARE:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    TOPOLOGY:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       US-07-667-276A-4
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           US-08-923-992A-4
                                                                                                                                                                                                                                                                                                                  COUNTRY:
                                                                                                                                                                                                                                                      STREET:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           LENGTH:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         Query Match
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            Matches
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        Ωy
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              qq
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       ŏλ
```

```
29;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       473 GSLDLMPLIPTIRETWTRKGITQKQHYSESADPTAISGMKRRAHSDRVFDFPPCDPTMDL 532
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    505 KEEKQVPKTPEQKHSKTEQNVPQE---PKSNDKNQLQELIKSAQQ------546
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     418 KNYQRLSDSVK-ARLVLENDDVSWSVQDLLPLCQELNIPLVLDWHHHNIVPG----TLRE 472
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  264 ERVFCSRTCRITTIQRDGLESVKQLGTQNVLDLIKLVEWNHNFGIHFMRVSSDLFPFASH 323
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             350 ------HFRK-------377
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  324 AKYGYTLEFAQSHLEEVGKLANKYNHRLTMHPGQYTQIASPREVVVDSAIRDLAYHDEIL 383
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 384 SRMKLNEQLNKDAVLIIHLGGTFEGK------KETL-------DRFR 417
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                401 NQQKIELTVSPENITV-----YEGEDVKFTVTAKSDSKTTLDFSDLLTKYNPSVSDRIS 454
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     455 TNYKTNTDNHKIAEITIKN------LKLNQSQTVTLKAKDDSGNVVEKTFTITVQK 504
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            193 KKIEDIRKQAQQADKKEDAEV--KVREELGKLFSSTKAGLDQQIQEHVKKETSSEENTQK 250
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                210 RRRSRSSAKNLEKE----STMNLDDHAPREMFDCLDKPI--PWRGRLGYACLNTILRSMK 263
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                308 YKAMSESLEQVEKELKHNSEANLODLVAK-----SKEIVREYEGKLN-----349
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               159 - EEQISKATKKVKQKEEE-----EYVEEVDEKSLKNESSSDEFEPVVPEQLETPISK 209
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              251 VDEHYANSLONLAQKSLEELDKATTNEQATOVKNOFLENAQKLKEIOPLIK----ETNVKL 307
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             44 HCLPDTLKSLLPMSSKTTLSMLPQVNIGANSFSAETPVDLKKENETELANISGPHKKS-- 101
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        142 HNKPNQOKTLKQSDTKVDLS----NIDKELNHQKSQV----EAMAEQAGITNEDKDSML 192
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               102 -TSTSTRKRARSSKKKATDSVSDKIDESVASYDSSTHLRRSSRSKKPVNYNSSSESES-- 158
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         Query Match
4.2%; Score 130; DB 4; Length 1104;
Best Local Similarity 19.7%; Pred. No. 0.013;
Matches 119; Conservative 88; Mismatches 187; Indels 210; Gaps
Sterne, Kessler, Goldstein & Fox P.L.L.C.
                                                                                                                                                                                                                           SYSTEM: PC-DOS/MS-DOS
PatentIn Release #1.0, Version #1.30
                             1100 New York Avenue, N.W., Suite 600
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    NAME: Esmond, Robert W. REGISTRATION NUMBER: 32,893 REFERENCE/DOCKET NUMBER: 1438.0140001/RWE
                                                                                                                                                                                                                                                                                                APPLICATION NUMBER: US/08/923,992A FILING DATE: 05-SEP-1997 CLASSIFICATION: 536
                                                                                                                                                                                                                                                                                                                                                                     PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 60/024,707
FILING DATE: 06-SEP-1996
ATTORNEY/AGENT INFORMATION:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      TELECOMMUNICATION INFORMATION:
TELEPHONE: (202) 371-2600
TELEPAX: (202) 371-2540
INFORMATION FOR SEQ ID NO: 4:
                                                                                                                                                                                                         IBM PC compatible
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                SEQUENCE CHARACTERISTICS:
LENGTH: 1104 amino acids
                                                                                                                                                                         Floppy disk
                                                                                                                                                                                                                                                                                CURRENT APPLICATION DATA:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               , MOLECULE TYPE: protein US-08-923-992A-4
                                                                                                                                                      COMPUTER READABLE FORM:
                                                                                                                                                                                                                                   OPERATING SYSTEM:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  amino acid
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               linear
                                                    Washington
                                                                                                                                                                                   MEDIUM TYPE:
                                                                                                                              20005
                                                                                                                                                                                                           COMPUTER:
     ADDRESSEE:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            TOPOLOGY:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         QQ
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        Q
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            В
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         ö
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              δ
```

```
73 NSFSAETPVDLKKENETELANISGPHKKSTSTSTRKRARSSKKKATDSVSDKIDESVASY 132
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            243 SSHKEKRPVDARGDEKSSVMGREKSHKASSKEESRRLL-----SEDSAKEKLPSSVVKK 296
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       133 DSSTHLRRSSRSKKPVN--YNSSSESESEQISKATKKVKQKEEEEFYVEEVDEKSLKNES 190
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   297 EKD---REGNSLKKKLSPALDVASDNHFKRPKHKDSEKIKSDKNKQSVDSVD-----S 346
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             191 SSDEFEPV-----VPEQLETPISKRRS--RSSAKNLEKESTMNLDDHAPR-----EMF 237
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       :1: | | : | | : | | : | | : | | : | | 347 GRGTGDPLPRAKDKVPNNLKAQEGKVRTNSDRKSPGSLPKVEEMDMDDEFEQPTMSFESY 406
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         Query Match
Best Local Similarity 20.0%; Pred. No. 0.01;
Matches 88; Conservative 72; Mismatches 168; Indels 111; Gaps
                                       547 -- ELEKLEKAIKEL-----MEQPEIP---SNPEY------GIOKSIW----ESQKE 582
533 MIEAKEKEQAVFELCRRYELQNPPCPLEIMGPEYDQTRDGYYPPGAEKRLTARKRRSRKE 592
                                                                                                                                                                                                                                                                                                                     APPLICANT: CONAWAY, RONALD C.
APPLICANT: CONAWAY, JOAN W.
TITLE OF INVENTION: ELONGIN A AND C FUNCTIONAL DOMAINS
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    /note= "amino acids 1-552 and 557-773 of Elongin A"
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              _IBM PC_compatible
SYSTEM: PC-DOS/MS-DOS
PatentIn Release #1.0, Version #1.30
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                . ., Ve
....вк: US/08/725,459В
04-OCT-1996
7N: 530
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        REFERENCE/DOCKET NUMBER: 11146/07501
                                                                                                                                                                                                                                                                                                                                                                                                         CORRESPONDENCE ADDRESS:
ADDRESSEE: SIDLEY & AUSTIN
STREET: 717 N. HARWOOD, SUITE 3400
CITY: DALLAS
                                                                                                                                                                                                                                                 Sequence 37, Application US/08725459B Patent No. 6084068
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    ATTORNEY/AGENT INFORMATION:
NAME: HANSEN, EUGENIA S.
REGISTRATION NUMBER: 31,966
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             TELECOMMUNICATION INFORMATION: TELEPHONE: 214-981-3300
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      SSS: not relevant not relevant
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 37:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              769 amino acids
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              CURRENT APPLICATION DATA:
APPLICATION NUMBER: US,
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       SEQUENCE CHARACTERISTICS:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          TELEFAX: 214-981-3400 INFORMATION FOR SEQ ID NO:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  MOLECULE TYPE: protein
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       OTHER INFORMATION:
OTHER INFORMATION:
                                                                                                                                                                                                                                                                                                                                                                                         NUMBER OF SEQUENCES:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 NAME/KEY: Protein
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    1..769
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     amino acid
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      OPERATING SYSTEM:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            COUNTRY: US
ZIP: 75201-6507
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               FILING DATE: 04 CLASSIFICATION:
                                                                                                                                                                                                                                                                                                GENERAL INFORMATION:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        STRANDEDNESS:
                                                                                                593 EVEE 596
                                                                                                                                            583 PIQE 586
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      US-08-725-459B-37
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     COMPUTER:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  TOPOLOGY:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      LOCATION:
                                                                                                                                                                                                                                   US-08-725-459B-37
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                SOFTWARE:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                LENGTH:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       STATE:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           q
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               QQ
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       ò
                                                     qq
```

```
238 DCLDKPIPWRGRLGYACLNTILRSMKERVFCSRTCRITTIQRDGLESVKQLGT-QNVLDL 296
                                                                                   297 IKLVEWNHNFGIHFM------RVSSDLFPF-----ASHAKYG--YTLEFAQS-- 335
                                                                                                                            448 OKLPKANENKSDKLOPAGAEPTRPRKVPTDVLPALPDIPLPAIQTNYRPLPSLELISSFQ 507
                                                                                                                                                                                                          508 PKRKAFSSPQEEEEAGFTGRRMNSKMQVYSGSKCAYLPKMMTLHQVLKNNIDSIFEVGGV 567
                                                                                                                                                                                                                                                       368 -----VVDSAIRDLAYHDEILSRWKLNEQLNKDAVLIIHLGGTF-EGKKETLDRFRKN 419
                                           -----LTMHPGQYTQIASPREV--- 367
                                                                                                                                                                                                                                                                                         568 PYSVLEPVLERCTPDQLYRIEECNHVLIEE---TDQLWKVHCHRDFKEERPEEYESWREM 624
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               4.1%; Score 128.5; DB 3; Length 769; 20.0%; Pred. No. 0.01; tive 72; Mismatches 168; Indels 111; Gaps
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       APPLICANT: CONAWAY, RONALD C.
APPLICANT: CONAWAY, JOAN W.
TITLE OF INVENTION: ELONGIN A AND C FUNCTIONAL DOMAINS
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             /note= "amino acid 1-556 and 561-773 of Elongin A"
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            SOFTWARE: Patentin Release #1.0, Version #1.30 CURRENT APPLICATION DATA:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             ATIONNE, COLOR HANSEN, EUGENIA S.
NAME: HANSEN, EUGENIA S.
RGISTRATION NUMBER: 31,966
REFERENCE/DOCKET NUMBER: 11146/07501
TELECOMMUNICATION INFORMATION:
TELEPHONE: 214-981-3400
TELEFAX: 214-981-3400
TELEFAX: 214-981-3400
INFORMATION FOR SEQ ID NO: 38:
SEQUENCE CHARACTERISTICS:
LENGTH: 769 amino acids
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    CORRESPONDENCE ADDRESS:
ADDRESSEE: SIDLEY & AUSTIN
STREET: 717 N. HARWOOD, SUITE 3400
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         APPLICATION NUMBER: US/08/725,459B
FILING DATE: 04-OCT-1996
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          Sequence 38, Application US/08725459B Patent No. 6084068 GENERAL INFORMATION:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
                                                                                                                                                                       336 -----HLEEVGKLANKYNHR---
                                                                                                                                                                                                                                                                                                                                           420 YORLSDSVKARLVLENDDV 438
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              SS: not relevant not relevant
                                                                                                                                                                                                                                                                                                                                                                                    625 YLRLQDAREQRLRLLTNNI 643
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         ATTORNEY/AGENT INFORMATION:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 Local Similarity 20.0% hes 88; Conservative
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             COMPUTER READABLE FORM:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         protein
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     TITLE OF INVENTION: HUMBER OF SEQUENCES:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    LOCATION: 1..769
OTHER INFORMATION:
OTHER INFORMATION:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                NAME/KEY: Protein
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    COUNTRY: US
ZIP: 75201-6507
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            FILING DATE: 04 CLASSIFICATION:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         DALLAS
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       STRANDEDNESS:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 MEDIUM TYPE:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         MOLECULE TYPE:
                                           407 LSYDQP-
                                                                                                                                                                                                                                                                                                                                                                                                                                                                   US-08-725-459B-38
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  US-08-725-459B-38
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     TOPOLOGY:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 STATE:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 Query Match
  δy
                                           g
                                                                                      δ
                                                                                                                            qq
                                                                                                                                                                                                              ρp
                                                                                                                                                                                                                                                       ŏ
                                                                                                                                                                                                                                                                                                Ωp
                                                                                                                                                                     ŏ
```

73 NSFSAETPVDLKKENETELANISGPHKKSTSTSTRKRARSSKKKATDSVSDKIDESVASY 132

Matches

```
133 DSSTHLRRSSRSKKPVN--YNSSSESEEQISKATKKVKQKEEEEYVEEVDEKSLKNES 190
                                                                                             191 SSDEFEPV-----VPEQLETPISKRRS--RSSAKNLEKESTMNLDDHAPR-----EMF 237
                                                                                                                                                                                                 347 GRGTGDPLPRAKDKVPNNLKAQEGKVRTNSDRKSPGSLPKVEEMDMDDEFEQPTMSFESY 406
                                                                                                                                                                                                                                             238 DCLDKPIPNRGRLGYACLNTILRSMKERVFCSRTCR1TTIQRDGLESVKQLGT-QNVLDL 296
                                                                                                                                                                                                                                                                                         407 LSYDQP-----RKKKKKVVKTSG---TALGEKGLKKKDSKSTSKNLNSA 447
                                                                                                                                                                                                                                                                                                                                     297 IKLVEWNHNFGIHFM------RVSSDLFPF-----ASHAKYG--YTLEFAQS-- 335
                                                                                                                                                                                                                                                                                                                                                                   | | : | | : | | : | | : | 448 OKLPRANENKSDKLQPAGAEPTRPRKVPTDVLPALPDIPLPAIQTNYRPLPSLELISSFQ 507
                                                                                                                                                                                                                                                                                                                                                                                                                                                                            508 PKRKAFSSPQEEEEAGFTGRRMNSKMQVYSGSKCAYLPKMMTLHQQCIRNIDSIFEVGGV 567
                                                                                                                                                                                                                                                                                                                                                                                                                                          ----- LTMHPGQYTQIASPREV--- 367
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        368 -----VVDSAIRDLAYHDEILSRMKINEQLNKDAVLIIHLGGTF-EGKKETLDRFRKN 419
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             568 PYSVLEPVLERCTPDQLYRIEBCNHVLIEE---TDQLWKVHCHRDFKEERPEEYESWREM 624
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 Patent No. booses.

GENERAL INFORMATION:
APPLICANT: CONAMAX, RONALD C.
APPLICANT: CONAWAX, JOAN W.
APPLICANT: CONAWAY, JOAN W.
TOWNENTION: ELONGIN A AND C FUNCTIONAL DOMAINS
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
                                                                                                                                                                                                                                                                                                                                                                                                                                336 -----HLEEVGKLANKYNHR-----
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   ADDRESSEE: SIDLEY & AUSTIN
STREET: 717 N. HARWOOD, SUITE 3400
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              APPLICATION NUMBER: US/08/725,459B
FILING DATE: 04-OCT-1996
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             Sequence 35, Application US/08725459B Patent No. 6084068
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC COMPATIBLE
OPERATING SYSTEM: PC-DOS/MS-
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      31,966
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     REFERENCE/DOCKET NUMBER: 11
TELECOMMUNICATION INFORMATION:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   420 YQRLSDSVKARLVLENDDV 438
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             SS: not relevant
not relevant
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               625 YLRLQDAREQRERLLTNNI 643
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             INFORMATION FOR SEQ ID NO: 35:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     ATTORNEY/AGENT INFORMATION:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           NAME: HANSEN, EUGENIA S. REGISTRATION NUMBER: 31.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      214-981-3300
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            LENGTH: 769 amino acids
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        SHOUENCE CHARACTERISTICS:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         214-981-3400
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          COMPUTER READABLE FORM:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      MOLECULE TYPE: protein
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  NAME/KEY: Protein
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      amino acid
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 75201-6507
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    FILING DATE: 04 CLASSIFICATION:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 DALLAS
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        STRANDEDNESS:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                Sn
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     US-08-725-459B-35
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      TELEPHONE:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                TOPOLOGY:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             COUNTRY:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       TELEFAX:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       STATE:
                                                                                                    qq
                                                                                                                                                       ó
```

```
362 -ASPREVVVDSAIRDLAYHDEILSRMKLNEQLNKDAVLIIHLGGTF-EGKKETLDRFRKN 419
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        336 ------HLEEVGKLANKYNHRLTMHPGQ---YTQI--------361
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     508 PKRKAFSSPQEEEEAGFTGRRMNSKMQVYSGSKCAYLTLHQQCIRVLKNNIDSIFEVGGV 567
                                                                                                                       Query Match
4.1%; Score 127.5; DB 3; Length 769;
Best Local Similarity 19.6%; Pred. No. 0.012;
Matches 86; Conservative 74; Mismatches 168; Indels 111; Gaps
                                                                                                                                                                                                                                                                                                                                                                                              191 SSDEFEPV-----VPEQLETPISKRRS--RSSAKNLEKESTMNLDDHAPR-----EMF 237
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       238 DCLDKPIPWRGRLGYACLNTILRSMKERVFCSRTCRITTIQRDGLESVKQLGT-QNVLDL 296
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      297 IKLVEWNHNFGIHFM------RVSSDLFPF-----ASHAKYG--YTLEFAQS-- 335
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              568 PYSVLEPVLERCTPDQLYRIEECNHVLIEE---TDQLWKVHCHRDFKBERPEEYESWREM 624
                                                                                                                                                                                                                73 NSFSAETPVDLKKENETELANISGPHKKSTSTSTRKRARSSKKKATDSVSDKIDESVASY 132
                                                                                                                                                                                                                                            243 SSHKEKRPVDARGDEKSSVWGREKSHRASSKEESRRLL-----SEDSAKEKLPSSVVKK 296
                                                                                                                                                                                                                                                                                                        133 DSSTHLRRSSRSKKPVN--YNSSSESESERQISKATKKVKQKEEEETVEEVDEKSLKNES 190
                                                                                                                                                                                                                                                                                                                                                      297 EKD---REGNSLKKKLSPALDVASDNHFKKPKHKDSEKIKSDKNKQSVDSVD----S 346
                                                                                                                                                                                                                                                                                                                                                                                                                                            347 GRGTGDPLPRAKDKVPNNLKAQEGKVRTNSDRKSPGSLPKVEEMDMDDEFEQPTMSFESY 406
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    GENERAL INFORMATION:
APPLICANT: CONMANY, RONALD C.
APPLICANT: CONMANY, JOAN W.
TITLE OF INVENTION: ELONGIN A AND C FUNCTIONAL DOMAINS NUMBER OF SEQUENCES: 79
              OTHER INFORMATION: /note= "amino acids 1-544 and OTHER INFORMATION: 549-773 of Elongin A"
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            OPERATING SYSTEM: PC_DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       11146/07501
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     UMBER: US/08/725,459B
04-0CT-1996
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       ADDRESSEE: SIDLEY & AUSTIN
STREET: 717 N. HARWOOD, SUITE 3400
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    ; Sequence 36, Application US/08725459B
; Patent No. 6084068
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               ATTORNEY/AGENT INFORMATION:
NAME: HANSEN, EUGENIA S.
REGISTRATION NUMBER: 31,96
REFERENCE/DOCKET NUMBER: 111
TELECOMMUNICATION:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                420 YORLSDSVKARLVLENDDV 438
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           625 YLRLQDAREQRERELTINI 643
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                COMPUTER READABLE FORM:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    CORRESPONDENCE ADDRESS:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   APPLICATION NUMBER:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     75201-6507
LOCATION: 1..769
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 CLASSIFICATION:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       DALLAS
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           FILING DATE:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       S
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        US-08-725-459B-36
                                                                  US-08-725-459B-35
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  STATE:
                                                                                                                                   Query Match
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          q
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      q
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        οy
                                                                                                                                                                                                                                                                           g
                                                                                                                                                                                                                                                                                                                                                                 g
                                                                                                                                                                                                                                                                                                                                                                                                                 ô
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  g
```

```
336 ------HLEEVGKLANKYNHRLTMHPG------QYTQIA------362
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     363 --SPREVVVDSAIRDLAYHDEILSRMKENEQENKDAVLIIHLGGTF-EGKKETLDRFRKN 419
                                                                                                                                                                                                                                                                                                                                                                                                Query Match
Best Local Similarity 19.6%; Pred. No. 0.012;
Matches 86; Conservative 75; Mismatches 167; Indels 111; Gaps
Matches 86; Conservative 75; Mismatches 167; Indels 111; Gaps
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             | : | | : | | : | | : | 448 QKLPRANENKSDKLQPAGAEPTRPRKVPTDVLPALPD1PLPAIOTNYRPLPSLELISSFQ 507
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        508 PKRKAFSSPQEEEEAGFTGRRMNSKMOVYSGSKCAYLPKMMOCIRVLKNNIDSIFEVGGV 567
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               568 PYSVLEPVLERCTPDOLYRIEECNHVLIEE---TDOLWKVHCHRDFKEERPEEYESWREM 624
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   238 DCLDKPIPWRGRLGYACLNTILRSMKERVFCSRTCRITTIQRDGLESVKQLGT-QNVLDL 296
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   407 LSYDQP-----RKKKKKKVVKTSG---TALGEKGLKKKDSKSTSKNLNSA 447
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               297 IKLVEWNHNFGIHFM------RVSSDLFPF-----ASHAKYG--YTLEFAQS-- 335
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  73 NSFSAETPVDLKKENETELAN1SGPHKKSTSTTRKRARSSKKKATDSVSDKIDESVASY 132
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             243 SSHKEKRPVDARGDEKSSVMGREKSHKASSKEESRRLL----SEDSAKEKLPSSVVKK 296
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            133 DSSTHLRRSSRSKKPVN--YNSSSESESEEQISKATKKVKQKEEEEEYVEEVDEKSLKNES 190
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           297 EKD----REGNSLKKKLSPALDVASDNHFKKPKHKDSEKIKSUKNKQSVDSVD-----S 346
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          191 SSDEFEPV-----VPEQLETPISKRRRS--RSSAKNLEKESTMNLDDHAPR-----EMF 237
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    347 GRGTGDPLPRAKDKVPNNLKAQEGKVRTNSDRKSPGSLPKVEEMDMDDEFEQPTMSFESY 406
                                                                                                                                                                                                                                                                                                                                                                                      h
Similarity 19.6%; Pred. No. 0.012;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                APPLICANT: CONAWAY, RONALD C.
APPLICANT: CONAWAY, JOAN W.
TITLE OF INVENTION: ELONGIN A AND C FUNCTIONAL DOMAINS
NUMBER OF SEQUENCES: 79
CORRESPONDENCE ADDRESS:
                                                                                                                                                                                                                                                                            OTHER INFORMATION: /note= "amino acids 1-548 and OTHER INFORMATION: 553-773 of Elongin A"
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 ADDRESSEE: SIDLEY & MUSICAL STREET: 717 N. HARWOOD, SUITE 3400
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    Sequence 39, Application US/08725459B Patent No. 6084068 GENERAL INFORMATION:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           IBM PC compatible
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          420 YQRLSDSVKARLVLENDDV 438
                                                                                                                                   SS: not relevant not relevant
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       625 YLRLQDAREORIRLLTNNI 643
                        TELEFAX: 214-981-3400
INFORMATION FOR SEQ ID NO: 36:
SEQUENCE CHARACTERISTICS:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     MEDIUM TYPE: Floppy disk
                                                                                           769 amino acids
214-981-3300
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               COMPUTER READABLE FORM:
                                                                                                                                                                                     MOLECULE TYPE: protein
                                                                                                                                                                                                                                      NAME/KEY: Protein
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  Y: US
75201-6507
                                                                                                             TYPE: amino acid
STRANDEDNESS: not
                                                                                                                                                                                                                                                             1..769
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     DALLAS
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            STATE: TX
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        US-08-725-459B-39
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              COMPUTER:
    TELEPHONE:
                                                                                                                                                                                                                                                                                                                                US-08-725-459B-36
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    COUNTRY:
                                                                                                                                                                   TOPOLOGY:
                                                                                                                                                                                                                                                             LOCATION:
                                                                                                LENGTH:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     CITY:
                                                                                                                                                                                                                FEATURE:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         qq
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                δ
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            qq
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        qq
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       οy
```

```
4.1%; Score 127.5; DB 3; Length 769;
19.6%; Pred. No. 0.012;
tive 75; Mismatches 167; Indels 111; Gaps
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              133 DSSTHLRRSSRSKKPVN--YNSSSESESEQISKATKKVKQKEEEETVEEVDEKSLKNES 190
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               73 NSFSAETPVDLKKENETELANISGPHKKSTSTSTRKRARSSKKKATDSVSDKIDESVASY 132
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       243 SSHKERRPVDARGDEKSSVMGREKSHKASSKEESRRLL----SEDSAKEKLPSSVVKK 296
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        297 EKD---REGNSLKKKLSPALDVASDNHFKKPKHKDSEKIKSDKNKQSVDSVD----S 346
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             191 SSDEFEPV-----VPEQLETPISKRRS--RSSAKNLEKESTMNLDDHAPR-----EMF 237
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       238 DCLDKPIPWRGRLGYACLNTILRSMKERVFCSRTCRITTIQRDGLESVKQLGT-ONVLDL 296
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   407 LSYDQP------RKKKKKVVKTSG---TALGEKGLKKKDSKSTSKNLNSA 447
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            297 IKLVEWNHNFGIHFM------RVSSDLFPF-----ASHAKYG--YTLEFAQS-- 335
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           336 -----HLEEVGKLANKYNHRLTMHPG------QYTQIA----- 362
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 448 OKLPKANENKSDKLQPAGAEPTRPRKVPTDVLPALPDIPLPAIQTNYRPLPSLELISSFQ 507
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                508 PKRKAFSSPQEEEEAGFTGRRMNSKMOVYSGSKCAYLPKMMTLHQQCIRVLKNIFEVGGV 567
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            363 --SPREVVVDSAIRDLAYHDEILSRMKLNEQLNKDAVLIIHLGGTF-EGKKETLDRFRKN 419
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          568 PYSVLEPVLERCTPDQLYRIEECNHVLIEE---TDQLWKVHCHRDFKEERPEEYESMREM 624
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          ELONGIN A AND C FUNCTIONAL DOMAINS
                                                                                                                                                                                                                                                                                                                                                                                                                                                                               /note= "amino acids 1-560 and 565-773 of Elongin A"
                 SOFTWARE: Patentin Release #1.0, Version #1.30 CURRENT APPLICATION DATA:
                                                                                                                                                                                              11146/07501
                                                                   UMBER: US/08/725,459B
04-OCT-1996
        PC-DOS/MS-DOS
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      Sequence 40, Application US/08725459B Patent No. 6084068
                                                                                                                     ATTORNEY AGENT INFORMATION:
NAME: HANSEN, EUGENIA S.
REGISTRATION NUMBER: 31,966
                                                                                                                                                                                     REFERENCE/DOCKET NUMBER: 111
TELECOMOUNICATION INFORMATION:
TELEPHONE: 214-981-3300
TELEFAX: 214-981-3400
                                                                                                                                                                                                                                           TELEFAX: 214-981-3400
INFORMATION FOR SEQ ID NO: 39;
SEQUENCE CHARACTERISTICS:
                                                                                                                                                                                                                                                                                                                                                           5S: not relevant
not relevant
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         GENERAL INFORMATION:
APPLICANT: CONAWAY, RONALD C.
APPLICANT: CONAWAY, JOAN W.
TITLE OF INVENTION: ELONGIN A
                                                                                                                                                                                                                                                                                                                 : 769 amino acids amino acid
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            420 YQRLSDSVKARLVLENDDV 438
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  625 YLRLQDAREQRLRLLTNNI 643
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                Best Local Similarity 19.68
Matches 86; Conservative
                                                                                                                                                                                                                                                                                                                                                                                         MOLECULE TYPE: protein
                                                                                                                                                                                                                                                                                                                                                                                                                                     NAME/KEY: Protein LOCATION: 1..769
OTHER INFORMATION: OTHER INFORMATION:
                                                             APPLICATION NUMBER:
OPERATING SYSTEM:
                                                                                                    CLASSIFICATION:
                                                                                       FILING DATE:
                                                                                                                                                                                                                                                                                                                                                      STRANDEDNESS:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 US-08-725-459B-39
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                US-08-725-459B-40
                                                                                                                                                                                                                                                                                                                 LENGTH:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   Query Match
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               δλ
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              óγ
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     qq
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               δ
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  Db
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              δλ
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     а
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   q
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            δ
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  q
```

```
4.1%; Score 127.5; DB 3; Length 769;
19.6%; Pred. No. 0.012;
Live 75; Mismatches 167; Indels 111; Gaps
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                73 NSFSAETPVDLKKENETELANISGPHKKSTSTSTRKKRARSSKKKATDSVSDKIDESVASY 132
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    243 SSHKEKRPVDARGDEKSSVMGREKSHKASSKEESRRLL----SEDSAKEKLPSSVVKK 296
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        133 DSSTHLRRSSRSKKPVN--YNSSSESESGISKATKKVKOKEEEETVEEVDEKSLKNES 190
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            297 EKD----REGNSLKKKLSPALDVASDNHFKKPKHKDSEKIKSDKNKQSVDSVD----S 346
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  191 SSDEFEPV------VPEQLETPISKRRS--RSSAKNLEKESTMNLDDHAPR-----EMF 237
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        238 DCLDKPIPWRGRLGYACLNTILRSMKERVFCSRTCRITTIQRDGLESVKQLGT-QNVLDL 296
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          407 LSYDQP------RKKKKKVVKTSG---TALGERGLKKKDSKSTSKNLNSA 447
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  297 IKLVEWNHNFGIHFM------RVSSDLFPF-----ASHAKYG--YTLEFAQS-- 335
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        448 QKLPKANENKSDKLQPAGAEPTRPRKVPTDVLPALPDIPLPAIQTNYRPLPSLELISSFQ 507
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            336 -----QYTQIA-----362
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      363 --SPREVVVDSATRDLAYHDEILSRMKLNEQLNKDAVLIIHLGGTF-EGKKETLDRFRKN 419
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              508 PKRKAFSSPQEEEEAGFTGRRMNSKMQVYSGSKCAYLPKMMTLHQQCIRVLKNNIDSGGV 567
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          /note= "amino acids 1-564 and 569-773 of Elongin A"
                                                                                                                                                                                                                          SOFTWARE: Patentin Release #1.0, Version #1.30 CURRENT APPLICATION DATA:
                                                                                                                                                                                                                                                                                                                                                                                      11146/07501
                                          3: SIDLEY & AUSTIN
717 N. HARWOOD, SUITE 3400
                                                                                                                                                                                                                                                                    US/08/725,459B
                                                                                                                                                                                                              PC-DOS/MS-DOS
                                                                                                                                                                                      IBM PC compatible
                                                                                                                                                                                                                                                                                                                                                                            REFERENCE/DOCKET NUMBER: 1111
TELECOMMUNICATION INFORMATION:
TELEPHONE: 214-981-3300
                                                                                                                                                                                                                                                                                                                                                                31,966
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      SS: not relevant
not relevant
                                                                                                                                                                                                                                                                                                                                                                                                                                      TELEFAX: 214-981-3400
INFORMATION FOR SEQ LD NO: 40:
                                                                                                                                                                                                                                                                                04-OCT-1996
V: 530
                                                                                                                                                                          Floppy disk
                                                                                                                                                                                                                                                                                                                          ATTORNEY/AGENT INFORMATION:
NAME: HANSEN, EUGENIA S.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              420 YQRLSDSVKARLVLENDDV 438
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                769 amino acids
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              86; Conservative
                                                                                                                                                                                                                                                                                                                                                                                                                                                                              SEQUENCE CHARACTERISTICS:
                                                                                                                                                    COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy of
                      CORRESPONDENCE ADDRESS:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  MOLECULE TYPE: protein
                                                                                                                                                                                                                                                                                                                                                              REGISTRATION NUMBER:
                                                                                                                                                                                                                                                                  APPLICATION NUMBER:
NUMBER OF SEQUENCES:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  OTHER INFORMATION:
                                                                                                                                                                                                            OPERATING SYSTEM:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        NAME/KEY: Protein
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                LOCATION: 1..769
OTHER INFORMATION:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    _
=
=
=
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  TYPE: amino acid
STRANDEDNESS: not
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      Best Local Similarity
                                                                                                                                75201-6507
                                                                                                                                                                                                                                                                                                  CLASSIFICATION:
                                                                          DALLAS
                                                                                                                  Sn
                                                                                                                                                                                                                                                                                    FILING DATE:
                                      ADDRESSEE:
                                                                                                  Ϋ́
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  US-08-725-459B-40
                                                                                                                                                                                        COMPUTER:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    TOPOLOGY:
                                                        STREET:
                                                                                                                  COUNTRY:
                                                                            CITY:
STATE:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                LENGTH:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          Query Match
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              Matches
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            qq
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          Qλ
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            οy
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      Q
Q
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      š
```

음

```
155 ESESEEQISKATK----KVKQKEEEEYVEEV----DEKSLKNESSSDEFEPVVP-EQLET 205
                                                                                      361 IASPREVVVDSAIRDLAYHDEILSRMKLNEQLNKDAVLIIHLGGTFEGKKETLDRFRKNY 420
312 RVSSDLF-----PFAS------HAKYGYTLEFAQSHLEEVGKLANKYNHRLTMHPGQYTQ 360
                                          717 KVKSSAFKDVYIPVASITEEKRNGQSVYKITAKAEKLQQ--ELENKYVDNFSFY----- 768
                                                                                                                                                                                                                                                                                                                                                                                                   APPLICANT: Weiser, Jeffrey N.
APPLICANT: Plant, Andrew G.
TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR TREATMENT TITLE OF INVENTION: OF STREPTOCOCCUS PNEUMONIAE INFECTION
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      PANITCH SCHWARZE JACOBS & NADEL, P.C.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               SOFTWARE: Patentin Release #1.0, Version #1.30 CURRENT APPLICATION DATA:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    79; Mismatches 180;
                                                                                                                                                                                   421 QRLSDSVKARLVLENDDVSWSVQDL-LPLCQELN 453
                                                                                                                                                                                                                           819 --IKDTFTGRLIGEKDGKIYAIYNLKKPLFENLS 850
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          ADDRESSEE: PANITCH SCHWAREL 36th Floor STREET; 1601 Market Street, 36th Floor
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     CLASSIFICATION: 424
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 60/026,939
FILING DATE: 23-SEP-1996
ATTORNEY/AGENT INFORMATION:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         омыек: US/08/790,912
29-JAN-1997
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     COMPUTER: IBM PC compatible OPERATING SYSTEM: PC-DOS/MS-DOS
                                                                                                                                                                                                                                                                                                                                      Sequence 2, Application US/08790912
Patent No. 5976542
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        REFERENCE/DOCKET NUMBER: 76: TELECOMMUNICATION INFORMATION:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        REGISTRATION NUMBER: 36,317
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       (215) 567-2020
(215) 567-2991
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      TELEX: 831-494
INFORMATION FOR SEQ ID NO: 2:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       2052 amino acids
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         86; Conservative
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  SEQUENCE CHARACTERISTICS:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                10
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               , MOLECULE TYPE: protein US-08-790-912-2
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     Leary, Kathryn
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        CORRESPONDENCE ADDRESS:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                Pennsylvania
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                APPLICATION NUMBER:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         Philadelphia
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     NUMBER OF SEQUENCES:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              amino acid
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   linear
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        19103-2398
                                                                                                                                                                                                                                                                                                                                                                                     GENERAL INFORMATION:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   USA
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     FILING DATE:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            TELEPHONE:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 TOPOLOGY:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   TELEFAX:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   COUNTRY:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         LENGIH:
                                                                                                                                                                                                                                                                                                                      US-08-790-912-2
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             Matches
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      Op
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    δλ
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          οy
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    qq
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                ò
                                                                                                                                                qq
                                                                                                                                                                                             δy
                                                                                                                                                                                                                                qq
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       18;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         662 ----ENSGNTTSENGQTEPEKKLELRNVSDIELXSQTNGTYRQHVSLDGIPENTDTYFV 716
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    206 PISK-RRSRSSAKNLEKESTMNLDDHAPREMFDCLDKPIPWRGRLGYACLNTILRSMKE 264
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              623 NSGKIANENTGEVSNKPSDSKPPVEESNQPEKNGTATKP-----------------661
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               459 PDTVVS--DKGEPEQVAPLPEYKGNIEQVKPETPVEKTKEQGPEKTEEVPVKPTEETPVN 516
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       97 PHKKSTSTSTRKRARSSKKKATDSV--SDKIDESVASYDSSTHLRRSSRSKKPVNYNSSS 154
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               517 PNEGTTECTSIQEAENPVOPAEESTTNSEKVSPDTSSENTGE------VSSN 562
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                155 ESESEEQISKATK----KVKQKEEEEYVEEV----DEKSLKNESSSDEFEPVVP-EQLET 205
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           Gaps
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 47 PDTLKSLLPMSSKTTLSMLPQVNIGANSFSAETPVDLKKE---NETELANIS-----G 96
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        Query Match
Best Local Similarity 18.9%; Pred. No. 0.053;
Matches 86; Conservative 79; Mismatches 180; Indels 109;
                                                                                                                                                                          APPLICANT: Wester, Jeffrey N.
APPLICANT: Wester, Jeffrey N.
APPLICANT: Plaut, Andrew G.
TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR TREATMENT
TITLE OF INVENTION: OF STREPTOCOCCUS PNEUMONIAE INFECTION
NUMBER OF SEQUENCES: 10
                                                                                                                                                                                                                                                                                                        ADDRESSEE: PANITCH SCHWARZE JACOBS & NADEL, P.C. STREET: 1601 Market Street, 36th Floor
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            SOFTWARE: PatentIn Release #1.0, Version #1.30 CURRENT APPLICATION DATA:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               265 RVFCSRTCRITTIQRDGLESVKQLGTQNVLDLIKLVEWN-
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 APPLICATION NUMBER: US 60/026,939
FILING DATE: 23-SEP-1996
ATTORNEY/AGENT INFORMATION:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           APPLICATION NUMBER: US/08/790,912
FILING DATE: 29-JAN-1997
CLASSIFICATION: 424
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  7600-4U1
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              PC-DOS/MS-DOS
                                                                                                           Sequence 3, Application US/08790912
Patent No. 5976542
                                                                                                                                                                                                                                                                                                                                                                                                                                                                    E: Floppy disk
IBM PC compatible
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             36,317
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             REFERENCE/DOCKET NUMBER: 76
TELECOMMUNICATION INFORMATION:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         (215) 567-2020
(215) 567-2991
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             INFORMATION FOR SEQ ID NO: 3:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  SEQUENCE CHARACTERISTICS:
LENGTH: 1964 amino acids
  625 YLRLQDAREQRLRLLTNNI 643
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     NAME: Leary, Kathryn
REGISTRATION NUMBER:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        TOPOLOGY: linear MOLECULE TYPE: protein
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               PRIOR APPLICATION DATA:
                                                                                                                                                                                                                                                                     NUMBER OF SEQUENCES: 1
                                                                                                                                                                                                                                                                                                                                                                                                                                               COMPUTER READABLE FORM:
                                                                                                                                                                                                                                                                                                                                                                                Pennsylvania
                                                                                                                                                                                                                                                                                                                                                         Philadelphia
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 OPERATING SYSTEM:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 amino acid
                                                                                                                                                                                                                                                                                                                                                                                                                   19103-2398
                                                                                                                                                          GENERAL INFORMATION:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        831-494
                                                                                                                                                                                                                                                                                                                                                                                                                                                                       MEDIUM TYPE:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 STRANDEDNESS
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                TELEPHONE:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              COMPUTER:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  US-08-790-912-3
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   TELEFAX:
                                                                                                                                                                                                                                                                                                                                                                                                   COUNTRY:
                                                                                           US-08-790-912-3
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          TELEX:
                                                                                                                                                                                                                                                                                                                                                                                   STATE:
```

q

qq ò q δ

```
18;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 590 PNEGTTEGTSIQEAENPVQPAEESTTNSEKVSPDTSSENTGE------VSSN 635
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               532 PDTVVS--DKGEPPGVAPLPEYKGNIEQVKPETPVEKTKEQGPEKTEEVPVKPTEETPVN 589
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     97 PHKKSTSTSTRKRARSSKKKATDSV--SDKIDESVASYDSSTHLRRSSRSKKPVNYNSSS 154
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    Indels 109; Gaps
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   47 PDTLKSLLPMSSKTTLSMLPQVNIGANSFSAETPVDLKKE---NETELANIS-----G 96
769 -----LDKKAKE--ENTNFTSFSNLVKAINQNPSGTYHLAASLNANEVELGPDERSY 818
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              Query Match 4.1%; Score 127.5; DB 2; Length 2052: Best Local Similarity 18.9%; Pred. No. 0.057;
```

Db 735ENSGNTISENGQTEPEKKLELRNVSDIELYSQTNGTYRQHVSLDGIPENTDTYFV 789
OY 206 PISK-RRRSRSSAKNIEKESTMNLDDHAPREMFDCLDKPIPWRGRLGYACLNTILRSMKE 264 DD 696 NSGRIANENTGEVSNKPSDSKPPVEESNQPEKNGTATKP734 OY 265 RVFCSRTCRITIORDGLESVKQLGTQNVLDLIKLVEWNHNFGIHFM 311
696
265
OY 312 RVSSDLFPFASHAKYGYTLEFAQSHLEEVGKLANKYNHRLTMHPGOYTQ 360
312 790 361
312 790 361 842
312 790 361 842 421
312 790 361 842 421 892

Search completed: January 15, 2002, 13:51:01 Job time: 330 sec

•

```
GenCore version 4.5
Copyright (c) 1993 - 2000 Compugen Ltd.
```

OM protein - protein search, using sw model

January 15, 2002, 14:07:00; Search time 45.07 Seconds (without alignments) 487.292 Million cell updates/sec Run on:

Title:

US-09-724-296-2_COPY_230_828 3112 1 MLRLLKRNIQISKRIVFTIL.........KRLTARKRRSRKEEVEEDEK 599 Perfect score:

Sequence:

BLOSUM62 Gapop 10.0 , Gapext 0.5 Scoring table:

100059 seqs, 36664827 residues Searched:

100059 Total number of hits satisfying chosen parameters:

Minimum DB seq length: 0 Maximum DB seq length: 20000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

SwissProt_39:* Database : Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES	Description	54	Q13439 homo sapien	schiz	P87178 schizosacch	P32532 human parai	schizo	Q04632 saccharomyc	Q02455 saccharomyc		human					P36309 apple stem		P35177 saccharomyc		٥	caenc		_				Q02952 homo sapien	-	P12839 rattus norv	Q9utk5 schizosacch	Q02224 homo sapien	P40631 tetrahymena	lw0 homo sap	Q01550 xenopus lae
	ID	YWJD_BACSU	GOG4_HUMAN	GAR2_SCHPO	YB1D_SCHPO	RRPP_PI1HE	IF38_SCHPO	YMH1_YEAST	MLP1_YEAST	NKCR_HUMAN	RRPP_PI1HD	MYH3_HUMAN	MSP1_PLAFK	MSP1_PLAFW	KEL1_YEAST	POLR_ASGVP	SEC3_YEAST	SPT7_YEAST	NSR1_YEAST	USO1_YEAST	YKI5_CAEEL	RBB1_HUMAN	MYSA_CAEEL	MYHD_HUMAN	BIR1_SCHPO	YAYB_SCHPO	AKAC_HUMAN	SEC9_YEAST	NFM_RAT	ALM1_SCHPO	CENE_HUMAN	MLH_TETTH	DIL2_HUMAN	TANA_XENLA
	DB	П	1	٦	П	-	П	-	-	Н	-	Н	Н	Н	_	-	7	-	Н	-	-	-	-	-	-	٦	Н	-	-	П	Н	-	Н	1
	Length	320	2230	200	726	568	918	607	1875	1462	268	1940	1630	1639	1164	2105	1336	1332	414	1790	952	1257	1969	1938	997	783	1781	651	845	72	9	633	4	1744
æ	Query		5.1	4.9	4.8	4.7	4.7	4.7	4.6	4.6	4.4	4.4	٠	4.4	4.3	4.3	4.3	4.3	4.2	4.2	4.2	4.2	4.2	4.1	4.1	•		4.1	4.1	4.1	4.1	4.1	4.1	4 . 1
	Score	360	160	154	150.5	147	145.5	145	144	142.5	138	138	136.5	136.5	135	134	133.5	132.5	132	132	131.5	130	129.5	129	128.5	128	128	127.5	127.5	127.5	127.5	127	127	127
	Result No.		7	e	4	ις	9	7	60	თ	10	11	12	13	14	15	16	17	18	19	20	21	22	23	24	25	56	27	28	53	30	31	32	33

394 KDAVLIIHLGGTFEGKKETLDRFRKNYQRLSDSVKARLVLENDDVSWSVQDLLPLCQELN 453

> Db ÓΫ qq Qy

454 IPLVLDWHHHNIVPGTLREGSLDLMPLIPTIRETWTRKGITQKQHYSESADPTAISGMKR 513

001613 xenopus lae 024308 pisum sativ 024106 bos taurus 003410 rattus norv 009560 caenorhabdi 066431 dugbe virus 009202 caenorhabdi 099051 drosophila 0990508 fundulus he 099323 drosophila P82094 homo sapien P23471 homo sapien	AA .	e) late) EBGION.	m group;	databases.	This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See http://www.isb-sib.ch/announce/or send an email to license@lisb-sib.ch).	48E002 CRC64;	DB 1; Length 320; 5-13; ss 109; Indels 28; Gaps 8;	TIQRDGLESVKQLGTQNVLDLIKLVEWNHNFGIHFMRVSSDLFPFASHAKYGYTLEFA 333	2 - Z
APX_XENLA TOP2_PEA ADFP_BOVIN SCP1_RAT YQY1_CAEEL RRPL_DUGBV PR2_CAEEL PIT_DROME VITI_FUNHE MYSN_DROME TYSN_DROME	ALIGNMENTS PRT; 320 AA	Created) Last sequence update) Last annotation update) PROTEIN IN ACDA 5'REGION	<pre>YWJD. Bacillus subtilis. Bacillus/Clostridium group Bacillus/Staphylococcus group; Bacillus.</pre>	Danchin A.; EMBL/GenBank/	This SWISS-PROT entry is copyright. It is posterior the Swiss Institute of Bioinformat the European Bloinformatics Institute. The use by non-profit institutions as long modified and this statement is not removed. The entities requires a license agreement (See or send an email to license@isb.sh.ch).	 lete proteome. MW; 2FD35A710648E002	Score 360; DB 1; 1 Pred. No. 9e-13; 61; Mismatches 109;	KLVEWNHNFGIHFN: : : :	GOYTQIASPREVVV NQFTLFTSPKESVJ
		Created) Last seq Last ann PROTEIN	grou	v., the	s col ituti tics tics stit ment ense	1; - 1; - 0).	00 00 00 00 00 00	LDLI	т : III SFHP
1420 1462 450 997 964 4036 1704 1004 2314	STANDARD;	32, 32, 40, KDA	tes; Ba	Fuente 95) to	ntry is s Inst nformal it ins states a lice to lice	CAA89865.1; CAA89865.1; 11309; ywjD. protein; Complete O AA; 36896 MW;	11.6%; .larity 29.8%; .Conservative (TIQRDGLESVKQLGTQNVL	KYNHRE : 1 :THQLRT:
	ST	(Rel. (Rel. (Rel.	ubtilis Tirmicu Taphylo	de la MAY-19	S-PROT entr the Swiss I dean Bioinfo non-profit and this st requires a n email to	32; CAB BG1130 BG1130	nilarit Conse	SLESVKÇ	SVGKLAN : : :CELVK
126.5 126.5 126.5 125.5 125.5 124.5 124.5 124.5	T 1 BACSU YWJD_BACSU	P45864; 01-NOV-1995 (01-NOV-1995 (20-AUG-2001 (HYPOTHETICAL	YWJD. Bacillus subtilis Bacteria: Firmicu Bacillus/Staphylo	SEQUENCE FROM N.A STRAIN=168; Glaser P., de la Submitted (MAY-19	This SWISS-PR between the the European use by non- modified and entities requ	EMBL, 249782; CAA89865.1; EMBL, 299123; CAB15748.1; SubtiList; BG11309; ywjD. Hypothetical protein; Comp SEQUENCE 320 AA; 36896	Query Match Best Local Similarity Matches 84; Conser	9 -	
33 33 33 33 33 33 34 44 54 54 54	7 1	P458 01-1 01-1 20-7 HYPC	YWJD Baci Bact Baci NCBI	SEQUENT SEQUEN	Thir Det the use mod ent	EMB EMB Sub Hypo	Query M Best Lo Matches	27	334 90
	RESULT YWJD_B ID Y	AC DT DT DE	000 00 00 00 00 00 00 00 00 00 00 00 00	RR RA RL	888888888	SO W R R R SO S	Q u Be Ma	VQ 4	DP DP

```
This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation-the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See http://www.isb-sib.ch/announce/or send an email to license@isb-sib.ch).
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 COILED COIL (POTENTIAL).
COILED COIL (POTENTIAL).
COLLED COIL (POTENTIAL).
COILED COIL (POTENTIAL).
COILED COIL (POTENTIAL).
TPYKGGNLYHTDVSLFGEPTEFEYLRKVLFEY -> HLTKV
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           GOLGI MEMBRANE.
ALTERNATIVE PRODUCTS: DIFFERENTS ISOFORMS MAY ARISE BY ALTERNATIVE
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   DISEASE: ANTIGEN IN THE AUTOIMMUNE DISEASE SJOEGREN'S SYNDROME AND
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            SUBCELLULAR LOCATION: ASSOCIATED WITH THE CYTOPLASMIC FACE OF THE
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   Thesis (1994), Instituto municipal de investigacion medica, Spain.
-!- FUNCTION: MAY PLAY A ROLE IN VESICULAR TRANSPORT FROM THE TRANS-
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              AICTIRMSHSLENLPNLSICEKCFLSI (IN ISOFORM
                                                                                       GOG4_HUMAN STANDARD;

20.31439; 013436; 013564;

20.40G-2001 (Rel. 40, Created)

20.40G-2001 (Rel. 40, Last sequence update)

20.40G-2001 (Rel. 40, Last sequence update)

GOLGI AUTOANTIGEN, GOLGIN SUBFAMILY A 4 (TRANS-GOLGI P230) (256 KDA GOLGIN) (GOLGIN-245) (72.1 PROTEIN).
                                                                                                                                                                                                                                                                                                                                               MEDLINE-96215236; PubMed-8626529; Erlich R., Gleeson P.A., Campbell P., Dietzsch E., Toh B.-H.; Molecular characterization of trans-Golgi p230: a human peripheral membrane protein encoded by a gene on chromosome 6p12-22 contains extensive coiled-coil alpha helical domains and a granin motif."; J. Biol. Chem. 271:8328-8337(1996).
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    MEDLINE-96125112; PubMed-8537393; Fritzler M.J., Chan E.K.L.; Fritzler M.J., Lung C.-C., Hamel J.C., Griffith K.J., Chan E.K.L.; Molecular characterization of golgin-245, a novel Golgi complex protein containing a granin signature."; Drotein Containing a Brain signature."; J. Biol. Chem. 270:31262-31268(1995).
                                                                                                                                                                                                                                                      Eukaryota, Metazoa, Chordata, Craniata, Vertebrata, Euteleostomi,
Mammalia, Eutheria, Primates, Catarrhini, Hominidae, Homo.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      coil; Alternative splicing.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              Submitted (NOV-1994) to the EMBL/GenBank/DDBJ databases
257 RSHADYVDANFLLPLLERFRQWGTNIDFMIEAKQKDKALLRL 298
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   SEQUENCE OF 131-2230 FROM N.A.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   SEQUENCE OF 524-672 FROM N.A.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  Solgi stack; Antigen; Coiled
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        EMBL; U41740; AAC50434.1; -.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            EMBL; X76942; CAA54261.1; -.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       CAA58041.1; -.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        MIM; 602509; -.
InterPro; IPR000237; GRIP.
Pfam; PF01465; GRIP; 1.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    1214
2152
2185
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   237
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       1011
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    TISSUE=Gastric fundus;
                                                                                                                                                                                                                                      Homo sapiens (Human)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     IN HEPATITIS B.
                                                                                                                                                                                                                                                                                                                              SEQUENCE FROM N.A.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                           SEQUENCE FROM N.A.
                                                                                                                                                                                                                                                                                           NCBI_TaxID=9606;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         TISSUE=Placenta;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    1033
1259
2154
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       EMBL; X82834;
EMBL; U31906;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    SPLICING.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      Balague C.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           VARSPLIC
                                                                                                                                                                                                                       GOLGA4
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 Seelig
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       DOMAIN
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          DOMAIN
                                                      RESULT 2
GOG4_HUMAN
QQ
```

```
27;
                                                                                                                                                                                                                                                                                                                                                                                                                                  1451 SRFTQHQNTVKELQIQLELKSKEAYEKDEQINLLKEELDQQNKRFDCLKGEMEDDKSKM- 1509
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      1557 IEHKELVOKLOHFOELGEEKDNRVKEAEEKILTLENOVYS-----MKAELETKKKELEH 1610
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 1636 KLAELKRKAEOKIAAIKKOLLSOMEEKEEOYKKGTESHLSELNTKLOEREREVHILEEKL 1695
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           1806 KKYSLIVAQHVEKEGGKNNIQAKQNLENVFDDVQKTLQEKELTCQILEQKIKELDSCLVR 1865
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     1696 KSVESSOSETLIVPRSAKNVAAYT---EQEEADSOGCVOKTYEEKISVLORNLTEKEKLL 1752
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      1753 QRVGQEKEETVSSHFEMRCQYQERLIKLEHARAKQHEDQSMIGHL------QEELEEKN 1805
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     156 SESEEQISKATKKVKQKEEEE-YVEEVDEK--SLKNESSSDEFEPVVPEQLETPISKRRR 212
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     273 RITTIQRDGLESV----KQL-----GTQNVLDLIKLVEWNHNFGIHFMR--- 312
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              313 --VSSD-----LFPFASHAKYGYTLEFAQSHLEEVGKLANKYNHRLTMHPGQYTQ---- 360
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     459 DWHHHNIVPGTLR---EGSLDLMPLIPTIRETWTRKGITQKQHYSESADPTAISGMKRRA 515
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          866 QKEVHRVEMEELTSKYEKLQALQQMDGRNKPTELLEENTEEKSKSHLVQPKLLSNMEAQ- 1924
                                                                                                                                                                                                                                                                                                                                                                                                              96 GPHKKSTSTSTRKRARSSKKKATDSVSDKIDESVASYDSSTHLRRSSRSKKPVNYNSSSE 155
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              213 SRSSAKNLEKESTMNLDDHAPREMFDCLDKPIPWRGRLGYACLNTILRSMKERVFCSRTC 272
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   361 -- IASPREVVVDSAIRDLAYHDEILSRMKLNE-QLNKDAVLIIHLGGTFEGKKETLDRFR 417
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            418 KNYQRL------SDSVKARLVLEN--DDVSWSVQDLLPLCQ------ELNIPLVL 458
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              516 HSDRVFDFPPCDPTMDLMIEAKEKEQAVFELCR-----RYELQNPPCPLEIMGPEYD 567
                                                                                                                                                                                                                                                                                               Matches 122; Conservative 109; Mismatches 233; Indels 164; Gaps
                                                                                                                                                                                                                                                                                                                                     41 SRFHCLPDTLKSL---LPMSSKTTLSMLPQVNIGANSFSAETP-VD-LKKENETELANIS 95
                                                     SWLRSSS (IN ISOFORM 4).
                                                                                                                                                                                                                                                            Score 160; DB 1; Length 2230; Pred. No. 0.46;
                                                                   Y -> H (IN REF. 3).

Y -> H (IN REF. 3).

T -> A (IN REF. 3).

T -> A (IN REF. 3).

T -> A (IN REF. 3).

K -> E (IN REF. 3).

K -> N (IN REF. 3).

K -> N (IN REF. 3).

W; 3BB733DB1EA86134 CRC64;
                 MISSING (IN ISOFORM 2).
MISSING (IN ISOFORM 3).
FTSPRSGIF -> SWLRSSS (II
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 P41891; 013707;
01-NOV-1995 (Rel. 32, Created)
0-MAY-2000 (Rel. 39, Last sequence update)
00-AUG-2001 (Rel. 40, Last annotation update)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     500 AA
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 Schizosaccharomyces pombe (Fission yeast).
                                                                                                                                                                                                                                                            5.1%; Score 160;
19.4%; Pred. No. 0
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        1973 QERE------EKIKQEQEDLE 1987
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       568 QTRDGYYPPGAEKRLTARKRKSRKEEVE 595
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     PRT;
                                                                                                                                                                                                       261139 MW;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     STANDARD;
                 22230
2109
2230
188
220
220
276
276
584
630
630
                                                                                                                                                                                                       2230 AA;
                                                                                                                                                                                                                                                                                Rest Local Similarity
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          PROTEIN GAR2.
GAR2 OR SPAC140.02.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 GAR2_SCHPO
               VARSPLIC
VARSPLIC
                                                   VARSPLIC
                                                                                                                                                                                                                                                                Query Match
                                                                                                                                                 CONFLICT
                                                                                                                                                                                                       SEQUENCE
                                                                                          CONFLICT
                                                                                                              CONFLICT
                                                                                                                                CONFLICT
                                                                                                                                                                   CONFLICT
                                                                                                                                                                                     CONFLICT
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               GAR2_SCHPO
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 RESULT
                                                                                                                                                                                                                                                                                                                                                                                                                                               q
ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      qq
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            _{\rm Q}
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             QQ
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    QQ
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             q
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   \delta
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        qq
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              δ
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             q
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     δ
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              ŏ
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 qq
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     ID
DT
DT
DT
DT
OS
```

```
Db
  ŏ
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               qq
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                qq
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    ò
                                                                                                                                                                                                                                                                                                                 This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See http://www.isb-sib.ch/announce/or send an email to license@isb-sib.ch).
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   9
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  168 KVKQ------KEEEEYVEEVDEKSLKNESSSDEFEPVVPEQLETPISKRRRSRSS 216
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              108 KRARSSKKKATDSVSDKIDESVASYDSSTHLRRSSRSKKPVNYNSSSESESEEQISKATK 167
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       48 DTLKSLLPMSSKTTLSMLPQVNIGANSFSAETPVDLKKENETELANISGPHKKSTSTSTR 107
                                                                                                                            "gar2 is a nucleolar protein from Schizosaccharomyces pombe required for 18S rRNA and 40S ribosomal subunit accumulation."; Nucleic Acids Res. 23:1912-1918(1995).
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           35 EAAKEIAKQSSKTDVS--PK------KSKKEAKRASSPEPSKKSV----
                                                                                                                                                                                                            Brown D., Churcher C.M., Barrell B.G., Rajandream M.A., Wood V.;
Submitted (SEP-1997) to the EMBL/GenBank/DDBJ databases.
-!- FUNCTION: HELPS THE ASSEMBLY OF PRE-RIBOSOMAL PARTICLES
                                                                                                       Gulli M.-P., Girard J.-P., Zabetakis D., Lapeyre B., Melese T.,
Caizergues-Ferrer M.;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      4.9%; Score 154; DB 1; Length 500; 27.9%; Pred. No. 0.15;
                                                                                                                                                                                                                                                                                                                                                                                                                                       EMBL; AL163191; CAB86413.1; -.

BESP; P09651; 14A1.

InterPro; IPR00076; rrm; 2.

SMART; SM00360; RRM; 2.

PROSITE; PS50102; RRM; 2.

PROSITE; PS00030; RRM; 2.

PROSITE; PS00030; RRM; 1.

Ribosome biogenesis; RNA-binding; Nuclear protein; Repeat;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   Indels
                                                                                                                                                                                                                                                               -!- SUBCELLULAR LOCATION: NUCLEAR; NUCLEOLUS.
-!- SIMILARITY: BELONGS TO THE GAR FAMILY.
-!- SIMILARITY: CONTAINS 2 RNA RECOGNITION MOTIFS (RRM).
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    -> P (IN REF. 1),
9D37FAD0C5161A0B CRC64;
Eukaryota; Fungi; Ascomycota; Schizosaccharomycetes; Schizosaccharomycetales; Schizosaccharomycetaceae;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          RNA-BINDING (RRM) 1.
RNA-BINDING (RRM) 2.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  903
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   Mismatches
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            GLU/SER-RICH
                                                                                         MEDLINE=95319932; PubMed=7596817;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                33;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 52987 MW;
                                                                                                                                                                                                                                                                                                                                                                                                                             EMBL; Z48166; CAA88179.1; -.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   Conservative
                                                                                                                                                                                                                                                       CONTAINING 18S RRNA
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     366
339
500 AA;
                           Schizosaccharomyces
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     217 AKNLEKESTM 226
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                SESEEEEVV 191
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    Similarity
                                                                                                                                                                                   SEQUENCE FROM N.A.
                                                               SEQUENCE FROM N.A.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                rRNA processing.
DOMAIN 81
                                        NCBI_TaxID=4896;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                53;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     DOMAIN
CONFLICT
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      Query Match
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               SEQUENCE
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     Best Local
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           DOMAIN
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                182
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               Matches
 δ
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           q
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              pp
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     a
```

PRT;

STANDARD;

YB1D_SCHPO

YB1D_SCHPO

RESULT

```
This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See http://www.isb-sib.ch/announce/or send an email to license@isb.ch).
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     282 SKALALDADIDAALTDKEGWFIOFYSSECDDCDDVSTAWYAMANRWRGKLNVAHINCAV- 340
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                133 HLNPDTDPDIPSDEDVLTDEDTEEVASIQPALSTSVSSLSLASTAMSKSASASEFSGSSV 192
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          261 SMKERVFCSRTCRITTIQ-----RDGLESVKQLGTQNVLDLIKLVEWNHNFGIHFM 311
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              341 -----SKRACKQYSIQYFPTFLFFKEEAF--VEYVGLPNEGDLVSFAEEAANFEIR-- 389
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              359 TQIASPREVVVDSAI------RDLAYHDEILSRMKLN--------EQLNKD 395
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                396 AVLIIHLGGTFEGKKET-----LDR----FRKNYQRLSDSVKARLVLENDDVSW 440
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     441 SVQDLLPLCQELNI-----PLVLDWHHHNIVPGTLREGSLDLMPLIPTIRETWTRKGI 493
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            566 ---DLEAIEAAKNFHVNGKPSPTRFAWVNGKFWAQWLRKFDIDIEATGPRV----IVYNA 618
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               95 SGPHKKSTSTSTRKRARSSKKKATDSVSDKIDESVASYDSSTHLRRSSRSKKPVNYNSSS 154
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           Hilbert H., Duesterhoeft A., Wood V., Rajandream M.A., Barrell B.G.; Submitted (MAY-1997) to the EMBL/GenBank/DDBJ databases.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      Gaps
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     448 RVVSFPKLIVVRDGIASYYPAKMAQDNKDYRRILGWMKNNWLPVLPELRTSNSKEIFNDE
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               44 HCLPDT---LKS----LLPMSSKTTLSMLPQVNIGANSFSAETPVDLKKENETEL--ANI
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                155 ESESEEQISKATKKVKOKOKEEEETVEEVDEKSLKNESSSDEFEPVVPEQLETPISKRRRSR
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           243 VQDAEFSIQTAPSFPKEKEEKENTEETEE-----SKKSINPT------GT
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  215 SSAKNLEKESTMNLDDHA-----PREMFDCLDKPIPW-----RGRLGYACLNTILR
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         312 RVSSDLFPFASHAKYG--YTLEFAQSHLEEVGKLANKYNHRLTMHPG------QY
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    390 --EVELLDTVNAEKNGDVFFLYFYDDDSAEYLNIIRKTGIQLLGHANLYLTTSQQIAKKY
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  Indels 149;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            Score 150.5; DB 1; Length 726; Pred. No. 0.36;
15-JUL-1998 (Rel. 36, Created)
15-JUL-1998 (Rel. 36, Last sequence update)
15-DEC-1998 (Rel. 37, Last annotation update)
HYPOTHETICAL 81.2 KDA PROTEIN C3D6.13C IN CHROMOSOME II.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          05D6AE65A4C9D6AF CRC64;
                                                                                                                                                                                        Schizosaccharomyces pombe (Fission yeast).
Eukaryota; Funqi; Ascomycota; Schizosaccharomycetes;
Schizosaccharomycetales; Schizosaccharomycetaceae;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  223;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               22.3%; Pred ....
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      POTENTIAL
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        Pfam. PF00085; thiored; 2.
Hypothetical protein; Transmembrane.
TRANSMEM 675 695
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          726 AA; 81237 MW;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            InterPro; IPR000063; Thiored.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     EMBL; Z95620; CAB09115.1; -.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           Best Local Similarity 22.3%
Matches 127; Conservative
                                                                                                                                                                                                                                                                                                              Schizosaccharomyces
                                                                                                                                                                                                                                                                                                                                                                                                                         SEQUENCE FROM N.A.
                                                                                                                                                                                                                                                                                                                                              NCBI_TaxID=4896
                                                                                                                                                         SPBC3D6.13C
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                Query Match
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 SEQUENCE
```

26;

```
IF38_SCHPO
                                                                                                                                                                                                  q
                                                                                                                                                                                                                                                          ŏ
                                ŏ
                                                                                                                                          ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                              This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See http://www.isb-sib.ch/announce/or send an email to license@isb-sib.ch).
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       62 LSMLPQVNIGANSFSAETPVDLKKENETELANISGPHKKSTSTSTRKRARSSKKKATDSV 12.1
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                122 SDKIDESVASYDSSTHLR-----RSSR---SKKPVNYNSSSE------SESEE 160
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     91 SRRTGES----KTETHARILDQQGIHRASRRGTSPNPLPENMGDERNTRIDEDSPNERRH 146
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             161 QISKATKKVKQ-----KEEEEFYVEEVDEKSLKNESSSDEFEPVVPEQLETPISKRRRSR 214
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 193 NNGRSMETSST-----HSTR-ITDVITNPSP------ELEEAVLQRKKRRP 231
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    275 TTIQRDGLESVKQLGTQNVLDLIKLVEW-NHNFGIHFMRVSSDLFPFASHAKYGYTLEFA 333
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 334 QSHLEEVGKLANKYNHRLTMHPGQYTQIASPREVVVDSAIRDLAYHDEILSRMKLNEQLN 393
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          Query Match 4.7%; Score 147; DB 1; Length 568; Best Local Similarity 21.2%; Pred. No. 0.41; Matches 114; Conservative 87; Mismatches 214; Indels 122; Gaps
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          215 SSAKNLEKESTMNLDDHAPREMFDCLDKPIPWRGRLGYACLNTILRSMKERVFCSRTCRI 274
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           36 LSYKP-TEIGEDRSWLHNIIDNPKENKSSCKSDDNNKDRAISTSTQDH-RSSEE---SGI 90
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  MEDLINE-92295573; PubMed-1318610;
Power U.F., Ryan K.W., Portner A.;
The P genes of human parainfluenza virus type 1 clinical isolates are polycistronic and microheterogeneous.";
Virology 189:340-343(1992).
- i FUNCTION: THIS PROTEIN IS PROBABLY A COMPONENT OF THE ACTIVE
                                                                                                                                                                                                                                              01-OCT-1993 (Rel. 27, Created)
01-OCT-1993 (Rel. 27, Last sequence update)
15-DEC-1998 (Rel. 37, Last annotation update)
RNA POLYMERASE ALPHA SUBUNIT (EC 2.7.7.48) (NUCLEOCAPSID
                                                                                                                                                                                                                                                                                                                                                                                                                   Human parainfluenza 1 virus (strain CI-14/83).
Viruses; ssRNA negative-strand viruses; Mononegavirales;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  Pfam; PF01806; Paramyxo_p; 1.
Transferase; RNA-directed RNA polymerase; Nucleocapsid;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          568 AA; 64586 MW; 8795D10C1CD3CCD7 CRC64;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         232 TTIKRNOTRSER---TON-SELHKSTSGDSSNLEDHNTKTSOKIPP--
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            POLYMERASE. IT MAY FUNCTION IN TEMPLATE BINDING.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                              Paramyxoviridae; Paramyxovirinae; Paramyxovirus
                                                                                                                                                                                                  568 AA.
494 TOKOHYSESADPTAISGMKRRAHSDRVFD 522
                                                    619 AQDIYWDETAKGTPIS----IEKDTVFD 642
                                                                                                                                                                                                  PRT;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             InterPro; IPR002693; Paramyxo_P.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           EMBL; M74080; AAA46830.1; -.
                                                                                                                                                                                                  STANDARD;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             SEQUENCE FROM N.A.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     NCBI_TaxID=31606;
                                                                                                                                                                                                                                                                                                                                                                 PHOSPHOPROTEIN).
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   Phosphorylation
                                                                                                                                                                                               RRPP_PI1HE
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       SEQUENCE
                                                                                                                                          RESULT 5
RRPP_PIIHE
ò
                                                                                                                                                                                                                      AC DDT THE REPORT OF THE THE REPORT OF THE T
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  q
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               a
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         a
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                δŏ
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             δ
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       Ω
```

```
between the Swiss Institute of Bioinformatics and the EMBL outstation—
the European Bioinformatics Institute. There are no restrictions on its
use by non-profit institutions as long as its content is in no way
modified and this statement is not removed. Usage by and for commercial
entities requires a license agreement (see http://www.isb-sib.ch/announce/
or send an email to license@isb-sib.ch).
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    This SWISS-PROT entry is copyright. It is produced through a collaboration
                                           329 --AKLLVSLGVMKSQHEFELSRRASHQFAKRMLKSANYKEMTFNLCGMLLSVEKSLGNKV 386
                                                                                                                                                   443 QDLLPLCQELNIPLVLDWHHHNIVPGTLREGSLDLMPLIPTIRETWTRKGITQKQHYSES 502
394 KDAVLIIHLG-----GTFEGKKETLDRFRK-----NYQRLSDSVKARLVLENDDVSWSV 442
                                                                                                                                                                                                                          387 EENRTLLKQIQEEIDSSRDLHKRFSEYQKEQNSLMMANLSTLHIITDRGGKT----GDP 441
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         METHIONYL-TRNAI AND MRNA (BY SIMILARITY).
-!- SUBUNIT: EIF-3 IS COMPOSED OF AT LEAST 10 DIFFERENT SUBUNITS (BY
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           MEDLINE-98162722; PubMed=9501991;
Yoshioka S., Kato K., Nakai K., Okayama H., Nojima H.;
"Identification of open reading frames in Schizosaccharomyces pombe
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              NA Res. 4:363-369(1997).
1- FUNCTION: BINDS TO THE 40S RIBOSOME AND PROMOTES THE BINDING OF
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    STRAIN=972;
Skelton J., Churcher C.M., Barrell B.G., Rajandream M.A., Wood V.,
Volckaert G., Aert R.;
                                                                                                                                                                                                                                                                                                      503 ADPTAISGMKRRAHSDRV----FDFPPCDPTMDLMIEAKEKEQAVFELCRRYELQNP 555
                                                                                                                                                                                                                                                                                                                                                                         442 SDTTRSPSVFTKGKDNKVKKTRF-----DPSMEALGGOEFKPDLIREDELREDIRNP 493
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           PROBABLE EUKARYOTIC TRANSLATION INITIATION FACTOR 3 93 KDA SUBUNIT
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           DB 1; Length 918;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        Submitted (APR-2000) to the EMBL/GenBank/DDBJ databases.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             Schizosaccharomyces pombe (Fission yeast).
Eukaryota; Fungi; Ascomycota; Schizosaccharomycetes;
Schizosaccharomycetales; Schizosaccharomycetaceae;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   30-MAY-2000 (Rel. 39, Created)
20-AUG-2001 (Rel. 40, Last sequence update)
20-AUG-2001 (Rel. 40, Last annotation update)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            -!- SIMILARITY: BELONGS TO THE EIF3S8 FAMILY.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       Ouery Match

4.7%; Score 145.5; E Best Local Similarity 23.5%; Pred No. 0.9; Matches 103; Conservative 57; Mismatches
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              TIF33 OR SPAC4A8.16C OR SPAC823.01C.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    PRT;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  EMBL; AL355013; CAB90146.1; -.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               SEQUENCE OF 571-918 FROM N.A.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          EMBL; Z98762; CAB11485.1; -. EMBL; D89140; BAA13802.1; -.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             InterPro; IPR000717; PCI.
P(am; PF01399; PCI; 1.
SMART; SM00088; PINT; 2.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    STANDARD;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                014164; P78791; Q9P6P4;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     Schizosaccharomyces.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     SEQUENCE FROM N.A.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 NCBI_TaxID=4896;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         SIMILARITY).
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       STRAIN-PR745;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    SCHPO
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              (EIF3 P93)
```

20;

Gaps

Indels 107;

Mismatches 172;

```
This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation-the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     entities requires a license agreement (see http://www.isb-sib.ch/announce/or send an email to license@isb-sib.ch).
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           325 --PEEIT-PADIFKYLRAIFEARGKKST-DRSEQIRLLEKLSTIAVIDYORL--RVKVAL 378
                                                                                                 237 F-DCLDKPIPWRGRLGYACLNTILRSMKERVFCSRTCRITTIQRDGLESVKQLGTQNVLD 295
                                                                                                                                                                                                                                                                                                           296 LIKLVEWNHNFG--IHFMRVSSDLFPFASHAKYGYTLEFAQSHLEEVGKLANKYNHRLT- 352
GANSFSAETPVDLKKENETELANISGPHKKSTSTSTRKRARSSKKKATDSVSDKIDESVA 130
                                                                          131 SYDSSTHLRRSSRSKKPVNYNSSSESEEQISKATKKVKQKEEEEYVEEVDEKSLKNES 190
                                                                                                                                                       --STMNLDDHAPREM 236
                                                                                                                                                                                      113 ESESEESESEE-ESDESERSGPSSFLKKPEKEEAKPAGLKFLRGESSEESSDEEEGRRV 171
                                                                                                                                                                                                                                                                                                                                                215 --KCKEAGRNPPPYIEFLSALDQKLESADKA-----FIKSLDAANGRAFNALKQRVRK 265
                                                                                                                                                                                                                                                                                                                                                                                        387 KLNEQLNKDAVLIIHLGGTFE--GKKETLDRFRK-------NYORLSDSVKARL 431
                      172 VKSAKDK----RYEEFISCMETIKNAMS-----SNNWIVVSNEFDHLNKVSQ-----
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       Eukaryota; Fungi; Ascomycota; Saccharomycotina; Saccharomycetes;
Saccharomycetales; Saccharomycetaceae; Saccharomyces.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  004632; 003641; Created)
01-NOV-1997 (Rel. 35, Last sequence update)
01-NOV-1997 (Rel. 40, Last stancotation update)
HYPOTHETICAL 69.8 KDA PROTEIN IN RPL6A-DAKI INTERGENIC REGION.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        Brown D., Bowman S., Barrell B.G., Rajandream M.A.;
Submitted (OCT-1994) to the EMBL/GenBank/DDBJ databases.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      STRAIN-5288C / AB972;
Bowman S., Churcher C., Barrell B.G., Rajandream M.A.;
Submitted (AUG-1994) to the EMBL/GenBank/DDBJ databases
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           69828 MW; 57E34E2DE36664B6 CRC64;
                                                                                                                                                     191 SSDEFEPWVPEQLETPISKRRRSRSSAKNLEKE---
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   PRT;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       Saccharomyces cerevisiae (Baker's
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            432 VLENDDVSWSVQDLLPLCQ 450
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 SEQUENCE OF 88-607 FROM N.A.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            EMBL; Z38114; CAA86249.1; -.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  379 LAVRFDINTGSGOYMPIDO 397
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              EMBL, 246373; CAA86507.1; -. SGD; SO004536; YML071C. Hypothetical protein SEQUENCE 607 AA; 69828 MW
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    SEQUENCE OF 1-113 FROM N.A.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     STANDARD;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     STRAIN=S288C / AB972;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  NCBI_TaxID=4932;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   YMH1_YEAST
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               YMH1_YEAST
                                                                                                                                                                                        q
                                                                                                                                                                                                                              õ
                                                                                                                                                                                                                                                                   q
                                                                                                                                                                                                                                                                                                         ò
                                                                                                                                                                                                                                                                                                                                              q
                                                                                                                                                                                                                                                                                                                                                                                                                        q
                                                                                                                                                                                                                                                                                                                                                                                                                                                                 ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            ŏ
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                qq
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      õ
                                                                                                                                                                                                                                                                                                                                                                                      ò
                                                                          ò
```

Length 607

DB 1;

4.78; Score 145;

Query Match

```
This SWISS-PROT entry is copyright. It is produced through a collaboration
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          between the Swiss Institute of Bioinformatics and the EMBL outstation—the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See http://www.isb-sib.ch/announce/or send an email to license@isb-sib.ch).
                                           18;
                                                                                                                                                                                                                                                                                                                                                                                             | :| | | | : | | :| | | | | : | 434 ENEIENETVNKTEDKAEKEEEVVIKONTKEENIKVEVTPEEPSKSIDNKAE---- 490
                                                                                                                                                                                                                          316 ESSNESLIEMMVKRKIEVLREHVYMSLNVFLKSFLYDTNDLEIPFPEELESTVLRINGTN 375
                                                                                                                                                                                                                                                                                                                                                                   159 EEQISKAT-KKVKQKEEEEYVEEVDEKSLKNESSSDEFE--PVVPEQLETPISKRRRSRS 215
                                                                                                                                                                                                                                                                                                                                                                                                                                                            SAKNLEKESTMNLDDHAPREMFDCLDKPIPWRGRLGYACLNTILRSMKERVFCSRTCRIT 275
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    276 TIQRD---GLESVKQLGTQNVLDL1----KLVEWNHNFGIHFMRVSSDLFPFASHAKYGY 328
                                                                                                                                    258 LVKLLSTNVSVNSLKKILQYLNSIPPFDGKTNKSLLSV--FLAMRYKFITDEIASYPLDV 315
                                                                                                                                                                                                                                                                          KKSTSTSTRKKARSSKKKATDSVSDKIDESVASYDSSTHLRRSSRSKKPVNYNSSSESES 158
                                                                                                                                                                                                                                                                                                                   376 EEKEIEEKEKETKKEEYQKQDSVANN-EEDVTENKSIEDVQEEVQCKVE-GEDDGAERKT 433
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    491 -----KEEEEINKVEVTPEEP----SKKIRTSKRENKIPTNAVMLQFVDK-----C-IT 534
                                           Gaps
                                                                                        1 MLRLLKRNIQIS--KRIVFTILKQKAFKG-NHPCVPSVCTITYSRFHCLPDTL----- 50
                                                                                                                                                                                 51 ----KSLLPMSSKTTLSMLPQ-----VNIGANSFSAET---PVDLKKENETELANISGPH 98
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  SEQUENCE FROM N.A. MEDLINE-94205265; PubMed=8154186; Bou G., Esteban P.F., Baladron V., Gonzalez G.A., Cantalejo J.G., Bou G., Esteban P.F., del Rey F., Ballesta J.P.G., Revuelta J.L.; Remacha M., Jimenez A., del Rey F., Ballesta J.P.G., Revuelta J.L.; "The complete sequence of a 15,820 bp segment of Saccharomyces cerevisiae chromosome XI contains the UBIZ and MPLI genes and three
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           -!- FUNCTION: MYOSIN-LIKE PROTEIN THAT IS PROBABLY INVOLVED IN DNA
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                535 YVLKDLTRGLNSIK-LSDSVCLQLVYCSFRI,CDLNRNYHHLFLKKINDTSLF----
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                Eukaryota; Fungi; Ascomycota; Saccharomycotina; Saccharomycetes; Saccharomycetales; Saccharomycetaceae; Saccharomyces.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            Koelling R., Nguyen T., Chen E.Y., Botstein D.; "A new yeast gene with a myosin-like heptad repeat structure."; Mol. Gen. Genet. 237:359-369(1993).
                                           64;
                     Pred. No. 0.57;
3; Mismatches 160; Indels
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      -!- SIMILARITY: SOME, TO THE TPR ONCOGENE.
-!- CAUTION: REF.2 MISQUOTES THE GENE NAME AS "MPL.1".
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               (Rel. 29, Last sequence update) (Rel. 40, Last annotation update)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              PRT; 1875 AA
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            Saccharomyces cerevisiae (Baker's yeast).
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      MEDLINE=93247549; PubMed=8483450;
21.68; Pic.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       01-0CT-1993 (Rel. 27, Created)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             329 TLEFAQSHLEEVGKLANKY 347
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           586 T'FEQLARAIDKRAELASKY 604
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             new open reading frames.";
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    MLP1 OR YKR095W OR YKR415.
                Best Local Similarity 21.69
Matches 82; Conservative
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    20-AUG-2001 (Rel. 40, Las MYOSIN-LIKE PROTEIN MLP1.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              STANDARD;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     Weast 9:1349-1354(1993)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           SEQUENCE FROM N.A.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                NCBI_TaxID=4932;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   STRAIN=S288C;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 01-JUN-1994
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            MLP1_YEAST
Q02455;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     REPAIR
                                                                                                                                                                                                                                                                                                                                                                                                                                                            216
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      MLP1_YEAST
                                                                                                                                                                                                                                                                        66
                                                                                                                                    qq
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  qq
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                ò
                                                                                                                                                                                                                          qq
                                                                                                                                                                                                                                                                                                                     a
                                                                                                                                                                                                                                                                                                                                                                                                            qq
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       Д
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    δλ
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                δŏ
                                                                                                                                                                                 ò
                                                                                                                                                                                                                                                                        ŏ
                                                                                                                                                                                                                                                                                                                                                                   ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                            ó
```

```
1566 MEQSGEIDVVLRKQLEAKVQEKQ -----KELENEYNKKLQEELKDVPHSSHISDDERDK 1619
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 1620 LRAEIESRLREEF---NNELQAIKKKSFDEGKQQAMMKTTLLERKLAKMESQLSETKQSA 1676
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  1677 ESPPKSVNNVQNPLLGLPRKIEENSNSPFNPLLSGEKLLKLNSKSSSGGFNPFTSPSPNK 1736
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  1737 HIONDNDKRESLANKIDPPIHLEPS--FNIPASRGLISSSSTLSTDINDEELISNNPAQK 1794
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          1795 DSSNRNVQSEEDT-----EKKKEGEPVKRGEAIEEQTKSNKRPIDEV-GELKNDE 1843
                                                                                                                                                                                                                                                                                                                                                                              ---SEEQISKATKKVKQKEEEEYVEEVDEKSLKN 188
                                                                                                                                                                                                                                                                                                                                                                                                                                             189 ESSSDEFEPVVPEQLETPISKRRRSRSSAKNLEKESTMNLDD------HAPREMFDC 239
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  ------GTQN-VLDLIKLVEWNHNFGIHFMRVSSDLFPFASHAKYG----YTLEFAQS 335
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   336 HLEEVG----KLANKYNHRLTMHPGQYTQIASPREVVVDSAIRDLAYHDEILS----- 384
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    385 -RMKLNEQLNKDAVLIIHLGGTFEGKK-----ETLDRFRKNYQRLSDSVKARLVLEN 435
                                                                                                                                                                                                                                                                                                           88 ETELANISGPHKKSTSTRKRARSSKKKATDSVSDKIDESVASYDSSTHLRRSSRSKKP 147
                                                                                                                                                                                                                                                                        71; Mismatches 164; Indels 102; Gaps
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   "A cyclophilin-related protein involved in the function of natural killer cells.";
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               240 LDKPIPWRGRLGYACLNTILRSMKERVF----CSRTCRITTIQRDGLESVKQL----
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              Eukaryota, Metazoa, Chordata, Craniata, Vertebrata, Euteleostomi,
Mammalia, Eutheria, Primates, Catarrhini, Hominidae, Homo.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       01-APR-1993 (Rel. 25, Created)
30-MAY-2000 (Rel. 39, Last sequence update)
20-AUG-2001 (Rel. 40, Last annotation update)
NK-TUMOR RECOGNITION PROTEIN (NATURAL-KILLER CELLS CYCLOPHILIN-RELATED PROTEIN) (NK-TR PROTEIN).
                                                                                                                                                                                                                                4.6%; Score 144; DB 1; Length 1875; 21.4%; Pred. No. 2.6;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           MEDLINE-93133824; PubMed=8421688;
Anderson S.K., Gallinger S., Roder J., Frey J., Young H.A.,
Ortaldo J.R.;
                                                                                                                                                                     -> A (IN RÈF. 1).
683A0D34C9066867 CRC64;
                                                                                                                  COILED COIL (POTENTIAL).
COILED COIL (POTENTIAL).
COILED COIL (POTENTIAL).
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             1462 AA.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             PRT;
                                                                                                                                                                                   1875 AA; 218455 MW;
         EMBL, L01992, AAA34783 1; -.
EMBL, X73541, CAA51948 1; -.
EMBL, Z28320, CAA82174 1; -.
PIR, S38173; S38173.
                                                                                                                                                                                                                                                                        92; Conservative
                                                                                                   Coiled coil; DNA repair.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             STANDARD;
                                                                                                                                  1678
1866
301
                                                                                                                    487
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  Homo sapiens (Human).
                                                                               S0001803; MLP1
                                                                                                                                                                                                                                                                                                                                                                            148 VNYNSSSESE----
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        1844 DDTTENINE 1852
                                                                                                                                                                                                                                                          Similarity
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      436 DDVSWSVQD 444
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                SEQUENCE FROM N.A.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               NCBI_TaxID=9606;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    rissue=Blood
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          NKCR_HUMAN
P30414:
                                                                                                                                                                                                                                      Query Match
                                                                                                                                                                     CONFLICT
                                                                                                                                                                                     SEQUENCE
                                                                                                                                                                                                                                                        Best Local
                                                                                                                                    DOMAIN
                                                                                                                                                      DOMAIN
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         RESULT 9
NKCR_HUMAN
                                                                                                                                                                                                                                                                        Matches
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  289
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              Dρ
ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                          QQ
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                ÓΫ
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                ó
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    δλ
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      δλ
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             δ
```

```
This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation-the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See http://www.isb-sib.ch/announce/or send an email to license@isb-sib.ch).
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 32;
                                                                                                                                -!- SUBCELLULAR LOCATION: ATTACHED TO THE MEMBRANE VIA 1TS N-TERMINUS.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               962 GSCSNSENNRGKPQKHKHGSKENLKREHTKKVKEKLK--------GKKDKKHKA 1007
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        622 GQKPWKPS----YERIQEMKAKTTHLLPIQSTYSLANIKETGSSSSYHKREKNSESDQS 676
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                -----107
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    677 TYSKYSDRSSESSPRSRSRSRSRSYSRSYTRSRSLASSHSRSPSSRSHSRNKYSDHS 736
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          ----KRARSSKKKATDSVSDKIDES------VASYDSSTH 137
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              37 QCSRSSSYTSISSDDGRRAKRRESSGKK--NSVSHKKHSSSSEKTLHSKYVKGRDRSSC 794
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                138 LRRSSRSKKPVNYNSSSESESEEQISKATKKVKQKEEEEFYVEEVDEKSLKNESSSDEFEP 197
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    795 VRKYSESRSSLDYSSDSEQSSVQ----ATQSAQEKEKQGQMERTHNKQEKNRG-----EE 845
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          --SSAKNLEKESTMNL--DDH- 231
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              846 KSKSERECPHSKRPTLKENLSDHLRNGSKPKRKNYAGSKWDSESNSERDVTKNSKNDSHP 905
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  232 -----REMKERVFCS--- 269
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      906 SSDKEECEATSDSESEVSEIHIKVKP----TTKSSTNTSLPDDNGAWKSSKORTSTSDSE 961
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          270 RTCRITTIQR------DGLESVKQLGTQNVLDLIKLVEWNHNFGIHFMRVSSDLFPFA 321
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   322 SHAKYGY----TLEFAQSHLEEVGKLANKYNHRLTMHPGQYTQIASPREVVVDSAIRDLA 377
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   Gaps
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   27 GNHPCVPSVCTITYSRFHCLPDTLKSLLPMSSKTTLSMLPQVNIGA-----NSFSAET 79
                                                                      Submitted (SEP-1999) to the EMBL/GenBank/DDBJ databases.
-!- FUNCTION: COMPONENT OF A PUTATIVE TUMOR-RECOGNITION COMPLEX.
INVOLVED IN THE FUNCTION OF NK CELLS.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               Mismatches 250; Indels 263;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             DB 1; Length 1462;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       D98All47763EF527 CRC64;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   Cyclosporin; Isomerase; Rotamase; Repeat; Transmembrane.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     ARG-SER TANDEM REPEAT-RICH
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   PPIASE, CYCLOPHILIN-TYPE. ARG/LYS-RICH (BASIC).
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         ARG/LYS-RICH (BASIC).
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           ARG/LYS-RICH (BASIC).
ARG/SER-RICH.
Proc. Natl. Acad. Sci. U.S.A. 90:542-546(1993).
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           4.6%; Score 142.5; I
20.2%; Pred. No. 2.3;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 ARG/SER-RICH.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 ARG/SER-RICH.
                                                                                                                                                                                                                                                                                                                                                                                                                                   InterPro; IPR002130; CSA_PPIASE.
Pfam; PF00160; pro_isomerase; 2.
PRNTNTS; PR00153; CSAPPISMRASE.
PROSITE; PS00170; CSA_PPIASE_1; 1.
PROSITE; PS50072; CSA_PPIASE_2; 1.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          198 VVPEQLETPISKRRRSR-------
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               20.2%; FIL.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       165676 MW;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              80 PVDLKKENETELA-------
                                                                                                                                                                                                                                                                                                                                                          AF184110; AAD56402.1; -.
                                                                                                                                                                                                                                                                                                                                        EMBL; L04288; AAA35734.2; -.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 Conservative
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         457
1010
                                                                                                                                                                                                                                                                                                                                                                              PIR; A47328; A47328.
HSSP; Q27450; 1A58.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         AA;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         Best Local Similarity
Matches 153; Conserv
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         421
970
194
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               654
1311
1462
                                                        Anderson S.K.;
                                                                                                                                                                                                                                                                                                                                                                                                                  MIM; 161565
                                      REVISIONS
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             Query Match
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       SEQUENCE
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     DOMAIN
DOMAIN
DOMAIN
DOMAIN
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               DOMAIN
DOMAIN
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     DOMAIN
    δλ
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    qq
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        ŏ
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            q
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                οy
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    qd
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            a
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  ΟŊ
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    q
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            δ
```

us-09-724-296-2_copy_230_828.rsp

```
SDKIDESVASYDSSTHLR-----RSSR---SKKPVNYNSSSESEEQISKATKKVK 170
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           SEQUENCE OF 774-1940 FROM N.A.
MEDLINE=90033298; PubMed=2806546;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   MEDLINE=89263803; PubMed=2726495;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  TISSUE=Skeletal muscle;
MEDLINE=90235862; PubMed=1691980;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           myosin heavy chains.";
Eur. J. Biochem. 189:55-65(1990)
                                                                                                                                                     ----HSTR-ITDVITNPSP----
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  P11055; Q15492; 01-10L-1989 (Rel. 11, Created)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  SEQUENCE OF 856-1940 FROM N.A.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  MYOSIN HEAVY CHAIN) (SMHCE).
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    FEBS Lett. 256:21-28(1989).
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              sapiens (Human).
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      SEQUENCE FROM N.A.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            Arnold H.H.;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  chain cDNA
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          MYH3_HUMAN
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                Ношо
    122
                                                              171
                                                                                           145
                                                                                                                                                     204
                                                                                                                                                                                                                                                                                                         404
                                 91
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          qq
                                                                                         a
                                                                                                                       Qγ
                                                                                                                                                     QC
                                                                                                                                                                                   Qγ
                                                                                                                                                                                                               qq
                                                                                                                                                                                                                                             Οy
                                                                                                                                                                                                                                                                         qq
                                                                                                                                                                                                                                                                                                         \overset{\circ}{\circ}
                                                                                                                                                                                                                                                                                                                                   q
                                                                                                                                                                                                                                                                                                                                                                    ò
                                                                                                                                                                                                                                                                                                                                                                                                 q
                                                                                                                                                                                                                                                                                                                                                                                                                                δy
                                                                                                                                                                                                                                                                                                                                                                                                                                                           qq
                                                              ð
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See http://www.isb-sib.ch/announce/
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   25;
      ---DDKQVTQESKEKKVSENNE----T 1050
                                                                                                                                                                                                                                                           1215 EKSQI------NLIDKKWKPLQGVGNLAAPNAATSSAVEVKVLTTVPEMKPQGLR 1263
                                                                                                                                                                                                  158 INIV----LKQDMATEHPQAEVVKQESSMSESKVLGEVGKQDSSSASLASAGESTGKKEVA 1214
                                                                             -----DLSGKHDTVTVSSDLDQFTKDDSKLSISPTALN 1097
                                                                                                          432 VLEN------DDVSWSVQDLL-------PLCQELNIPL---VLDWHH 462
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              62 LSMLPQVNIGANSFSAETPVDLKKENETELANISGPHKKSTSTSTRKRARSSKKKATDSV 121
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   Gaps
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  36 LSYKP-TEIGEDRSWLHSIIDNSKENKPSCKADDNNKDRAISTPTQDH-RSSEE---SGI 90
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           Power U.F., Ryan K.W., Portner A.;
The P genes of human parainfluenza virus type 1 clinical isolates
are polycistronic and microheterogeneous.";
                                               378 YHDEILSRMKLNEQLNKDAVLIIHLGGTFEGKKET-----LDRFRKNYQRLSDSVKARL
                                                                                                                                       1098 TEENVACLONIQHVEESVPNGVEDVLQTDDNMEICTPDRSSPAKVEETSPLGNARLDTPD
                                                                                                                                                                                                                                  516 HSDRVFDFPPCDPTMDLMIEAKEKE-QAVFELCRRYELQNPPCPLEIM------GPE
                                                                                                                                                                      463 HNIVPGTLREGSLDLMPLIPTIRETWTRK-----GITQKQHYSESADPTA--ISGMKRRA
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           VICTORY 189:340-343(1992).
-!- FUNCTION: THIS PROTEIN IS PROBABLY A COMPONENT OF THE ACTIVE POLYMERASE. IT MAY FUNCTION IN TEMPLATE BINDING.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              21.1%; Pred. No. 1.2;
Live 93; Mismatches 222; Indels 100;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  4.4%; Score 138; DB 1; Length 568;
                                                                                                                                                                                                                                                                                                                                                                                                                               01-0CT-1993 (Rel. 27, Created)
01-FB-1996 (Rel. 33, Last sequence update)
15-DEC-1998 (Rel. 37, Last annotation update)
RNA POLYMERASE ALPHA SUBUNIT (EC 2.7.7.48) (NUCLEOCAPSID
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     Viruses; ssRNA negative-strand viruses; Mononegavirales;
Paramyxoviridae; Paramyxovirinae; Paramyxovirus.
NCBI_TaxID=31607;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           RNA-directed RNA polymerase; Nucleocapsid;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      64644 MW; FACC2D694E9ACE0C CRC64;
                                                                                                                                                                                                                                                                                                            566 YDQTRDGYYPPGA---EKRLTAR-KRRSRKEEVEEDEK 599
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           Human parainfluenza 1 virus (strain CI-5/73).
                                                                                                                                                                                                                                                                                                                                                                                                     568 AA
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 or send an email to license@isb-sib.ch).
                                                                                                                                                                                                                                                                                                                                                                                                     PRT;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    SEQUENCE FROM N.A.
MEDLINE=92295573; Pubmed=1318610;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          InterPro; IPR002693; Paramyxo_P. Pfam; PF01806; Paramyxo_P; 1.
                  1008 PKRKQAFHWQPPLEFGEEEEI-
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               EMBL; M74082; AAA46838.1; -.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  Best Local Similarity 21.19
Matches 111; Conservative
                                                                                                                                                                                                                                                                                                                                                                                                     STANDARD;
                                                                              1051 IKDNILKTEKSSEE---
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       568 AA;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              PHOSPHOPROTEIN).
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        Phosphorylation
SEQUENCE 568
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           Transferase;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               PIR; C40234;
                                                                                                                                                                                                                                                                                                                                                                                                     RRPP_PI1HD
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     Query Match
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           Q
                                                                                                                                                                                                                                                                                                                                                                                                     a
                                                                              g
                                                                                                                                        g
                                                                                                                                                                        ò
                                                                                                                                                                                                    g
                                                                                                                                                                                                                                  õ
                                                                                                                                                                                                                                                             QQ
                                                                                                                                                                                                                                                                                            ð
                                                  ò
```

```
01-JUL-1999 (Rel. 11, Last sequence update)
20-AUG-2001 (Rel. 40, Last annotation update)
MYOSIN HEAVY CHAIN, FAST SKELETAL MUSCLE, EMBRYONIC (MUSCLE EMBRYONIC
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   "Identification of three developmentally controlled isoforms of human
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    ----GTFECKKETLDRFRK-----NYQRLSDSVKARLVLENDDVSWSVQDLLPLCQELN 453
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               338 IKSOHEFELSRSASHVFAKRMLKSANYKEMTFNLCGMLLSVEKSLENKVEENRTLLKQIQ 397
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                286 KQLGTQNVLDLIKLV-EWNHNFGIHFMRVSSDLFPFASHAKYGYTLEFAQSHLEEVGKLA 344
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  ---SKNEESVATOK 284
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        285 NNHNHRKTKYT---TNNANNNTKSLPTPEHDTTANEEGTSNTSVDEM----AKLLVSLGV 337
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  454 IPLVLDWHHHNIVPGTLREGSLDLMPLIPTIRETWTRKGITQKQHYSESADPTAISGMKR 513
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         398 EEIDSSRDLHKRFSEYQKEQNSLMMANLSTLHIITDRGGKT-----GDPSDTTRSPSVFT 452
226 MNLDDHAPREMFDCLDKPIPWRGRLGYACLNTILRSMKERVFCSRTCRITTIQRDGLESV
                                                                                                                                                                                                                                                                                                                                                                                             ------ELEDAVLQRNKRRPTTIKRSQTRSE
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          345 NKYNHRLTMHPGQYTQIASPREVVVDSAIRDLAYHDEILSRMKLNEQLNKDAVLIIHLG-
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
NCBI_TaxID=9606;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            Eller M.S., Stedman H.H., Sylvester J.E., Fertels S.H., Wu Q.-L., Raychowdhury M.K., Rubinstein N.A., Kelly A.M., Sarkar S.; Raychowdhury m.K., Rubinstein n.A., Kelly A.M., Sarkar S.; "Human embryonic myosin how chain cDNA. Interspecies sequence conservation of the myosin rod, chromosomal locus and isoform specific transcription of the gene.";
                                                                                                                  OKEEEEYVEEVDEKSLKNESSSDEFE-PVVPEQLE--TPISKRR--RSRSSAKNLEKEST
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               Eller M.S., Stedman H.H., Sylvester J.E., Fertels S.H., Rubinstein N.A., Kelly A.M., Sarkar S.; "Nocleotide sequence of full length human embryonic myosin heavy chain CDNA."
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                ., Buchberger-Seidl A., Braun T., Singh S., Goedde H.W.,
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     RAHSDRV----FDFPPCDPTMDLMIEAKEKEQAVFELCRRYELQNP 555
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        453 KGKDNKVKKTRF----DPSMEALGGQEFKPDLIREDELRDDIRNP 493
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  243 R---TQN-SELHKSTSEDSSNLEDHNTKTSPKVLP----
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     PRT; 1940 AA
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      Nucleic Acids Res. 17:3591-3592(1989).
```

```
This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMEL outstation the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See http://www.isb-sib.ch/announce/or send an email to license@isb-sib.ch).
                                                                                                                                                                                                                                                  DOMAIN: THE RODLIKE TAIL SEQUENCE IS HIGHLY REPETITIVE, SHOWING CYCLES OF A 28-RESIDUE REPEAT PATTERN COMPOSED OF 4 HEPTAPEPTIDES, CHARACTERISTIC FOR ALPHA-HELICAL COILED COILS.
                                                                                                                             SUBUNIT: MUSCLE MYOSIN IS A HEXAMERIC PROTEIN THAT CONSISTS OF 2 HEAVY CHAIN SUBUNITS (MHC), 2 ALKALI LIGHT CHAIN SUBUNITS (MLC) and 2 REGULATORY LIGHT CHAIN SUBUNITS (MLC-2).
SUBCELLULAR LOCATION: THICK FIND FOR THE MYOFIBRILS.
DEVELOPMENTAL STAGE: ABUNDANTLY PRESENT IN FETAL SKELETAL MUSCLE AND NOT PRESENT OR BARELY DETECTABLE IN HEART AND ADULT SKELETAL
                                                                                                                                                                                                                                                                                                 PTM: TWO CYSTEINE REBIDUES IN THE SI DOMAIN ARE SELECTIVELY ALKIATED AND ARE REGUIRED FOR MYOSIN ATPASE ACTIVITY.

MISCELLANEOUS: EACH MYOSIN HEAVY CHAIN CAN BE SPLIT INTO 1 LIGHT MEROMYOSIN (LAM) AND 1 HEAVY MEROMYOSIN (HAM). IT CAN LATER BE SPLIT FURTHER INTO 2 GLOBULAR SUBFRAGMENTS (S1) AND 1 ROD-SHAPED
                                              Karsch-Mizrachi I., Travis M., Blau H., Leinwand L.A.;
"Expression and DNA sequence analysis of a human embryonic skeletal
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      Myosin; Muscle protein; Coiled coil; Thick filament; Actin-binding; Calmodulin-binding; ATP-binding; Methylation; Alkylation;
                                                                                                                                                                                                                                                                                                                                                                                                       SIMILARITY: CONTAINS 1 MYOSIN-LIKE GLOBULAR HEAD DOMAIN. SIMILARITY: CONTAINS 1 1Q DOMAIN.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             METHYLATION (TRI-) (POTENTIAL).
ALKYLATION (SH-1).
ALKYLATION (SH-2).
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             A -> G (IN REF. 3).

KK -> QE (IN REF. 1 AND 2).

SR -> RA (IN REF. 3).

SR -> OT (IN REF. 2).

W; 43CA58C6A4BA1253 CRC64;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            COILED COIL (POTENTIAL).
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          MYOSIN HEAD-LIKE.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            ATP (POTENTIAL).
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             ACTIN-BINDING.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               ACTIN-BINDING
                                                                                muscle myosin heavy chain gene.";
Nucleic Acids Res. 17:6167-6179(1989).
-!- FUNCTION: MUSCLE CONTRACTION.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          Pfam; PF00612; IQ; I...
Pfam; PF00063; myosin_head; I.
Pfam; PF01576; Myosin_tail; I.
PRINTS; PR00193; MYOSINHEAVY.
ProDom; PD000355; myosin_head; I.
            SEQUENCE OF 856-1940 FROM N.A.
MEDLINE=89366648; PubMed=2771643;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               : IPR001609; myosin_head.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               InterPro; IPR002928; Myosin_tail
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          EMBL, X13988, CAA32167.1; -- EMBL, X13100; CAA31492.1; -- EMBL, X51593; CAA35942.1; -- EMBL, X15696; CAA33731.1; --
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  224035
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      MYSc; 1.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               InterPro; IPR000048; IQ.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      PROSITE; PS50096; IQ; 1.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          781
811
1933
186
678
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 1664
                                                                                                                                                                                                                                                                                                                                                                                       SUBFRAGMENT (S2)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             PIR, S04090, S04090.
HSSP; P13538; 2MYS.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                1940 AA;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      SMART; SM00015; IQ;
SMART; SM00242; MYS
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          Multigene family
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              608
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               1391
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               MIM; 160720;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   InterPro
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  SEQUENCE
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               CONFLICT
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           NP_BIND
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           MOD_RES
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               CONFLICT
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               CONFLICT
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 CONFLICT
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             MOD_RES
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            DOMAIN
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            DOMAIN
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              DOMAIN
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             DOMAIN
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          DOMAIN
```

4.4%; Score 138; DB 1; Length 1940;

Query Match

```
25;
                                                                                                                                                                                            997 KKALQEAHQQALDDLQAEEDKVNSLNKTKSKLEQQ-VEDL-ESSLEQE-----K 1043
                                                                                                                                                                                                                                                                                    ---K 1082
                                                                                                                                                                                                                                                                                                                                                                                                                                                     1128 ---ATRAKTEKORSDYARELEELSERLEEAGGVTS-------TOIELNKKREAEF 1172
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             220 KEKSEFKLEIDDLSSSMESVSKSKANLEKICRTLEDQL-----SEARGKNEEIQRSLS 1272
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              1273 ELTTOKSRLOTEAGELSRQLEEKESIVSQLS-----RSKOAF----TOOTEELKROLEE 1322
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                1323 ENKAKNALAHALQSSRHDC--DILREQYEEEQEG----KAELQRALSKANSEVAQWRTKY 1376
                                                               84 KKENETELANISGPHKKSTSTSTRKRARSSKKK-ATDSVSDKIDESVASYDSS-THLRRS 141
                                                                                                        941 KRKLEDECSEL ---- KKDIDDLELTLAKVEKEKHATENKVKNLTEELSGLDETIAKLTRE 996
                                                                                                                                                                                                                                                                                                                          262 MKERVFCSRTCRITTIQRDGLE---SVKQLGTQNVLDLIKLVEWNHNFGIHFMRVSSDLF 318
                                                                                                                                                                                                                                                                                                                                                                                                            319 PFASHAK-----YGYTLEFAQSHLEEVGKLANKYNHRLIMHPGQYTQIA--SPREVVV 369
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                370 DSAIRDL----AYHDEILSRMKINEQLNKDAVLIIHLGGTFEGKKETLDRFRKNYQRLSD 425
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      426 SVKARLVLENDDVSWSVQ-----DLLPLCQELNIPLVLDWHHHNIVPGTLREGSLDLM 478
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     479 PLIPTIRETWTRKGITQKQHYSESADPTALSGMKRRAHSDRVFDFPPCDPTMDLMIEAKE 538
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         539 KEQAVFELCRRYELQNPPCPLEIMGPEYDQTRDGYYPPGAEKRLTARKRRS----- 589
                                                                                                                                                   142 SRSKKPVNYNSSSESESEEQISKATKKVKQKEEEEYVEEVDEKSLKNESSSDEFEPVVPE 201
                                                                                                                                                                                                                                    202 QLETPISKRRRSRSSAKNLEKESTMNLDDHAPREMFDCLDKPIPWRGRLGYACLNTILRS 261
                        Gaps
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                42
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       -!- PTM: MEROZÓITE SURFACE ANTIGEN CONTAIN THE SEQUENCE OF 83 KDA, 4 KDA AND 19 KDA ANTIGENS WHICH ARE THE MAJOR SURFACE ANTIGENS OF
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    13-AUG-1987 (Rel. 05, Created)
1-FEB-1996 (Rel. 33, Last sequence update)
01-GCT-1996 (Rel. 34, Last annotation update)
MEROZOITE SURFACE PROTEIN 1 PRECURSOR (MEROZOITE SURFACE ANTIGENS)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   Pan W., Tolle R., Bujard H.;
Submitted (JUN-1995) to the EMBL/GenBank/DDBJ databases.
-!- SUBCELLULAR LOCATION: ATTACHED TO THE MEMBRANE BY A GPI-ANCHOR
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             "Polymorphism of the precursor for the major surface antigens of Plasmodium falciparum merozoites: studies at the genetic level."; EMBO J. 4:3823-3829(1985).
  Pred. No. 5.7;
97; Mismatches 206; Indels 144;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    Mackay M., Goman M., Bone N., Hyde J.E., Scaife J., Certa U.,
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              Alveolata; Apicomplexa; Haemosporida; Plasmodium
                                                                                                                                                                                                                                                                               044 KLRVDLERNKRKLEGDLKLAQESILDLENDKQQ-----LDERL----
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        Plasmodium falciparum (isolate K1 / Thailand).
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                PRT; 1630 AA.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       SEQUENCE FROM N.A.
MEDLINE=86136024; PubMed=3004972;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 REVISIONS, SEQUENCE FROM N.A.
19.68;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   1377 ETDAIQRTEELEEAKK 1392
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             Stunnenberg H., Bujard H.;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           590 ----RKEEVEEDEK 599
                      Conservative
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                STANDARD;
Best Local Similarity
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              NCBI_TaxID=5839;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             (PMMSA) (P190).
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       (POTENTIAL)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            MSP1_PLAFK
P04932;
                      109;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              Eukaryota;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     RESULT 12
                      Matches
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         MSP1_PLAFK
                                                                                                                                                                                                                                       QΫ
                                                                                                                                                                                                                                                                             Q
                                                                                                                                                                                                                                                                                                                          δλ
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    Db
                                                                                                                                                                                                                                                                                                                                                                                                            ^{\circ}
                                                                                                                                                                                                                                                                                                                                                                                                                                                   qq
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    Ωý
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          QQ
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   δλ
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              qq
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             Q
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                Qγ
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           Ω
```

```
This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation-the European Bioinformatics Institute. There are no restrictions on its mes by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See http://www.isb-sib.ch/announce/or send an email to license@isb-sib.ch).
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          35;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        407 EGKKETLDRFRKNYQ--RLSDSVKAR----LVLENDDVSWSVQDLLPLCQELNIPLVLDW 460
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  423 PFDYTKEPSKNIYTDNERKKFINEIKEKIKIEKKKIESDKKSYEDRSKSLNDITKEYEKL 482
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    176 FDLLRAKL---NDVCANDYCQIPFNLKIRANEL-DVLKKLV----------- 212
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  SFSAETPVDLKKENETELANISGPHKKSTSTRKRARSSKKKATDSVSDKIDESVASYD 133
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 213 -FGYRKPLDNIKDNVGKMEDYI---KKN------KKTIENINELIEESKKTID 255
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                134 SSTHLRRSSRSKKPVN-----YNSSSESESEEQISKATKKVKQKEEEEYVEEVDEK-- 184
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                -PISKRRRSRSSA 217
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               375 KNIDISAKVETKEST-----EPNEYPNGVTYP-----LSYNDINNALNELNSFGDLIN 422
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   -----KERVFCSRTCRITTIQRDGLESVKQL---GTQNVLDLIKLVE-- 301
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               -----WNHNFGIHFMRVSSDLFPFAS--HAKYGYTLEFAQSHLEEVGKLANKYNHRL 351
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          Indels 217; Gaps
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       17 FTILKQKAFKGNHPCVPSVCTITYS---RFHCLPDTLKSLLPMSSKTTLSMLPQVNIGAN 73
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            483 LNEIYDSKFNNNI------DLTNFEKMMGKRYSYKVEKLTHHNTFASYENSKHNLEK
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          352 TMHPGQYTQIASPREVVVDSAIRDLAYHDEILSRMK-----LNEQLNKDAVLITHLGGTF
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       KNLE-----KESTMNLDDHAPREMFDCLDKPIPWRGRLGYACLNTILRSM-----
                                                                                                                                                                                                                                                                                                                                                         (POTENTIAL)
                                                                                                                                                                                                                                                                                                                                                                        (POTENTIAL)
                                                                                                                                                                                                                                                                                                                                                                                      (POTENTIAL)
                                                                                                                                                                                                                                                                                                                                                                                                                                   (POTENTIAL)
                                                                                                                                                                                                                                                                                                                                         (POTENTIAL)
                                                                                                                                                                                                                                                                                                                                                                                                      (POTENTIAL)
                                                                                                                                                                                                                                                                                                                                                                                                                   (POTENTIAL)
                                                                                                                                                                                                                                                                                                                                                                                                                                                   (POTENTIAL)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                    (POTENTIAL)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 . .) (POTENTIAL)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 (POTENTIAL)
                                                                                                                                                                                                                 Malaria; Merozoite; Polyprotein; Repeat; Signal; Glycoprotein;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            DB 1; Length 1630;
MEROZOITES, THE MATURATION TAKE PLACE DURING SCHIZONT.
                                                                                                                                                                                                                                                             MEROZOITE SURFACE PROTEIN 1. TRIPEPTIDE SG(TP) REPEAT.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               ADBDEC3CE0A46322 CRC64;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           Best Local Similarity 19.5%; Pred. No. 5.5;
Matches 130; Conservative 107; Mismatches 212;
                                                                                                                                                                                                                                                                                                                          (GLCNAC. .
                                                                                                                                                                                                                                                                                                                                         (GLCNAC.
                                                                                                                                                                                                                                                                                                            (GLCNAC.
                                                                                                                                                                                                                                                                                                                                                                      (GLCNAC.
                                                                                                                                                                                                                                                                                        MEMBRANE ANCHOR.

N-LINKED GLCNAC.

                                                                                                                                                                                                                                                                                                                                                                                                    (GLCNAC.
                                                                                                                                                                                                                                                                                                                                                                                                                                     (GLCNAC
                                                                                                                                                                                                                                                                                                                                                                                                                                                     (GECNAC
                                                                                                                                                                                                                                                                                                                                                                                                                                                                    (GLCNAC
                                                                                                                                                                                                                                                                                                                                                                                                                                                                               (GLCNAC
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  (GLCNAC
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           --ESSSDEFEPVVPEQLET-
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           4.4%; Score 136.5; 19.5%; Pred. No. 5.5
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 N-LINKED
                                                                                                                                                                                                                                                                                                                                                                                                                                                                    N-LINKED
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 N-LINKED
                                                                                                                                                                    PIR; A25120; SAZQK1.
InterPro; IPR000561; EGF-like.
                                                                                                                                                    EMBL; X03371; CAA27070.1; -.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 1630 AA; 187289
                                                                                                                                                                                                                                   GPI-anchor
                                                                                                                                                                                                                                                                                                                                                                                                                     955
1049
1156
1165
1436
                                                                                                                                                                                                                                                1630
1630
84
1630
97
                                                                                                                                                                                                Pfam; PF00008; EGF; 1
                                                                                                                                                                                                                                  Transmembrane;
                                                                                                                                                                                                                                                                             DOMAIN
TRANSMEM
                                                                                                                                                                                                                                                                                                                                                                                        CARBOHYD
CARBOHYD
                                                                                                                                                                                                                                                                                                                                                                                                                                     CARBOHYD
                                                                                                                                                                                                                                                                                                                                                                                                                                                       CARBOHYD
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    CARBOHYD
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  CARBOHYD
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               SEQUENCE
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                Query Match
                                                                                                                                                                                                                                                                                                             CARBOHYD
                                                                                                                                                                                                                                                                                                                          CARBOHYD
                                                                                                                                                                                                                                                                                                                                            CARBOHYD
                                                                                                                                                                                                                                                                                                                                                           CARBOHYD
                                                                                                                                                                                                                                                                                                                                                                        CARBOHYD
                                                                                                                                                                                                                                                                                                                                                                                                                      CARBOHYD
                                                                                                                                                                                                                                                                                                                                                                                                                                                                    CARBOHYD
                                                                                                                                                                                                                                                SIGNAL
                                                                                                                                                                                                                                                                CHAIN
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       218
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     263
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                q
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         q
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 q
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              q
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       q
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          ò
ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          ò
```

```
This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation the European Bioinformatics Institute. There are no restrictions on its most by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See http://www.isb-sib.ch/announce/or send an email to license@isb-sib.ch).
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     Holder A.A., Lockyer M.J., Odink K.G., Sandhu J.S., Riveros-Moreno V., Nicholls S.C., Hillman Y., Davey L.S., Tizard M.L.V., Schwarz R.T.,
                                                                                                       503 ADPTA---ISGMKRRAHSDRVFDFPPCDPTMDLMIEAKEKEQAVFELCRRYELQNPPCPL 559
             634
                                                                                                                                    692 SEPSTEGEITGGATTKPGQQAGSALEGD---SVQAQAQEQKQA------QPPVPV 737
                                                                                                                                                                                                                                                                                                               ol-FEB-1996 (Rel. 33, Last sequence update)
01-FEB-1996 (Rel. 33, Last annotation update)
MEROZOTTE SURFACE PROTEIN 1. PRECURSOR (MEROZOTTE SURFACE ANTIGENS)
(PMMSA) (P195).
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               Submitted (MAR-1991) to the EMBL/GenBank/DDBJ databases.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          PTM: MEROZOITE SURFACE ANTIGEN CONTAIN THE SEQUENCE OF 83 KDA, KDA AND 19 KDA ANTIGENS WHICH ARE THE MAJOR SURFACE ANTIGENS OMEROZOITES. THE MATURATION TAKE PLACE DURING SCHIZONT.
--LD----LMPLIPTIRETWTRKGITQKQHYSES
                                                                         635 - KHNIHVPNSYKQENKQEPYYL1VLKKEIDKLKVFMPKVESLINEEKKNIKTEGQ--SDN
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           (POTENTIAL)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           (POTENTIAL)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            (POTENTIAL)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         (POTENTIAL)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               (POTENTIAL)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              (POTENTIAL) (POTENTIAL)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                "Primary structure of the precursor to the three major surface antigens of Plasmodium falciparum merozoites."; Nature 317:270-273(1985).
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  Pfan, PE00008; EGF: 1.
Malaria; Merozoite; Polyprotein; Repeat; Signal; Glycoprotein;
                                                                                                                                                                                                                                                                                                                                                                                                                            Eukaryota; Alveolata; Apicomplexa; Haemosporida; Plasmodium
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              PROTEIN 1.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           N-LINKED (GLCNAC.
N-LINKED (GLCNAC.
N-LINKED (GLCNAC.
N-LINKED (GLCNAC.
N-LINKED (GLCNAC.
N-LINKED (GLCNAC.
N-LINKED (GLCNAC.
N-LINKED (GLCNAC.
N-LINKED (GLCNAC.
N-LINKED (GLCNAC.)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              MEROZOITE SURFACE
                                                                                                                                                                                                                                                                                                                                                                                                             Plasmodium falciparum (isolate Wellcome).
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          MEDLINE=86014355; PubMed=2995820;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    InterPro; IPR000561; EGF-like.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     EMBL; X02919; CAA26676.1; -
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  GPI-anchor
                                                                                                                                                                                                                                                                                   STANDARD;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 1639
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          844
920
964
1058
                                             461 HHHNI-VPGTLREGS-
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    PIR; A24594; A24594.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                          SEQUENCE FROM N.A.
                                                                                                                                                                                                                                                                                                                                                                                                                                            NCBI_TaxID=5848;
                                                                                                                                                                                            : 11
738 PV--PE 741
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               (POTENTIAL)
                                                                                                                                                                       560 EIMGPE 565
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    Transmembrane;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        Freeman R.R.;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    Holder A.A.;
                                                                                                                                                                                                                                                                                   PLAFW
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   REVISIONS
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                CARBOHYD
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           CARBOHYD
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 CARBOHYD
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              CARBOHYD
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               CARBOHYD
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              CARBOHYD
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              CARBOHYD
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                CARBOHYD
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    SIGNAL
                                                                                                                                                                                                                                                                                   MSP1_PL
P04933;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  CHAIN
                                                                                                                                                                                                                                                                     MSP1_PLAFW
                                                                                                                                                                                                                                                    RESULT 13
                                                                                                                                                                                                                                                                                   Ob
                                                                                                          Óγ
                                                                                                                                        d
                                                                                                                                                                       òγ
                                            ò
```

(POTENTIAL)

CARBOHYD

```
35;
                                                                                                                                                                                                                                                                                                            74 SFSAETPVDLKKENETELANISGPHKKSTSTSTRKRARSSKKKATDSVSDKIDESVASYD 133
                                                                                                                                                                                                                                                                                                                                         222 -FGYRKPLDNIKDNVGKMEDYI---KKN-------KKTTENINELIEESKKTID 264
                                                                                                                                                                                                                                                                      185 FDLLRAKL---NDVCANDYCQIPFNLKIRANEL-DVLKKLV----------- 221
                                                                                                                                                                                                                                                                                                                                                                                              134 SSTHLRRSSRSKKPVN-----YNSSESEEEQISKATKKVKQKEEEEYVEEVDEK-- 184
                                                                                                                                                                                                                                                                                                                                                                                                                                     265 KNKNATKEEEKKKLYQAQYDLSIYNKQLE-EAHNLISVLEKRIDTLKKNENIKELLDKIN 323
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  185 SLKN-------PISKRRRSRSSA 217
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   324 EIKNPPPANSGNTPNTLLDKNKKIEEHEKEIKEIAKTIKFNIDSLFTDPLELEYYLREKN 383
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         384 KNIDISAKVETKEST -----EPNEYPNGVTYP-----LSYNDINNALNELNSFGDLIN 431
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               263 -----GTQNVLDLIKLVE-- 301
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            492 LNEIYDSKFNNNI------DLTNFEKMMGKRYSYKVEKLTHHNTFASYENSKHNLEK 542
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          EGKKETLDRFRKNYQ--RLSDSVKAR----LVLENDDVSWSVQDLLPLCQELNIPLVLDW 460
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        461 HHHNI-VPGTLREGS------LD----LMPLIPTIRETWTRKGITQKQHYSES 502
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           432 PFDYTKEPSKNIYTDNERKKFINEIKEKIKIEKKKIESDKKSYEDRSKSLNDITKEYEKL 491
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     302 -----WNHNFGIHFMRVSSDLFPFAS--HAKYGYTLEFAQSHLEEVGKLANKYNHRL, 351
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    352 TMHPGQYTQIASPREVVVDSAIRDLAYHDEILSRMK-----LNEQLNKDAVLIIHLGGTF 406
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              594 E-KKITKDENKPDEKILEVSDIVKVQVQKVLLMNKIDELKKTQLILK----NVEL--- 643
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                644 -KHNIHVPNSYKQENKQEPYYLIVLKKEIDKLKVFMPKVESLINEEKKNIKTEGQ--SDN 700
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         503 ADPTA---ISGMKRRAHSDRVFDFPPCDPTMDLMIEAKEKEQAVFELCRRYELQNPPCPL 559
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 701 SEPSTEGEITGOATTKPGQOAGSALEGD---SVQAQAQEQKQA------QPPVPV 746
                                                                                                                                                                                                                               17 FTILKQKAFKGNHPCVPSVCTITYS---RFHCLPDTLKSLLPMSSKTTLSMLPQVNIGAN 73
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 218 KNLE-----KESTMNLDDHAPREMFDCLDKPIPWRGRLGYACLNTILRSM-----
N-LINKED (GLCNAC. .) (POTENTIAL).
N-LINKED (GLCNAC. .) (POTENTIAL).
N-LINKED (GLCNAC. .) (POTENTIAL).
N-LINKED (GLCNAC. .) (POTENTIAL).
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    Saccharomyces cerevisiae (Baker's yeast).
Bukaryota; Fungi; Ascomycota; Saccharomycotina; Saccharomycetales;
Saccharomycetales; Saccharomycetaceae; Saccharomyces.
NCBI_TaxID=4932;
                                                                                                                                                                                           Indels 217;
                                                                                                                                               4.4%; Score 136.5; DB 1; Length 1639; 19.5%; Pred. No. 5.6;
                                                                                   2C255B6616C87F6E CRC64:
                                                                                                                                                                 Best Local Similarity 19.5%; Pred. No. 5.6;
Matches 130; Conservative 107; Mismatches 212;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  01-FEB-1995 (Rel. 31, Last sequence update)
20-AUG-2001 (Rel. 40, Last annotation update)
KELCH REPEATS PROTEIN 1.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             PRT; 1164 AA.
                                                           1526 1526 N-L
1639 AA; 187618 MW;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            01-FEB-1995 (Rel. 31, Created)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           STANDARD;
1165
1174
1445
1526
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                KEL1 OR YHR158C.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           560 EIMGPE 565
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 747 PV--PE 750
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    KEL1_YEAST
P38853;
  CARBOHYD
                   CARBOHYD
                                         CARBOHYD
                                                             CARBOHYD
                                                                                   SEQUENCE
                                                                                                                                               Query Match
FT
FT
FT
SO
                                                                                                                                                                                                                                                                    Dp
                                                                                                                                                                                                                                                                                                              ò
                                                                                                                                                                                                                                                                                                                                                      g
                                                                                                                                                                                                                                                                                                                                                                                                δy
                                                                                                                                                                                                                                                                                                                                                                                                                                     q
                                                                                                                                                                                                                                                                                                                                                                                                                                                                               δ
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       Dp
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    q
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   δ
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        Dp
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    δλ
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            qq
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         οD
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 Ωy
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    οy
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         δλ
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               qq
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           ŏ
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         AC DT DT OC OC OC OC
```

```
This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the BMBL outstation. The European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial conticies requires a license agreement (See http://www.isb-sib.ch/announce/or send an email to license@isb-sib.ch).
                                                                                                                                                                                                                                                                                                                                                               in cell fusion and morphology in Saccharomyces cerevisiae.";
J. Cell Biol. 143:375-389(1998).
-!- FUNCTION: HAS A ROLE IN CELL MORPHOGENESIS AND CELL FUSION AND MAY
                                                                     Johnston M., Andrews S., Brinkman R., Cooper J., Ding H., Dover J., Du Z., Favello A., Fulton L., Gattung S., Geisel C., Kirsten J., Kucaba T., Hillier L., Jier M., Johnston L., Langston Y., Latreille P., Louis E.J., Macri C., Mardis E., Menezes S., Mouser L., Nhan M., Rifkin L., Riles L., St Peter H., Trevaskis E., Vaughan K., Vignati D., Wilcox L., Wohldman P., Waterston R., Wilson R.,
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               62 LSMLPQVNIGANSFSAETPVDLKKENETELANISGPHKKSTSTSTRKRARSSKKKATDSV 121
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   616 LETLPSNE1-----KTP---QNGNIEEIKHLPDADEKIDSTTTFDQEINGDKLGTSSM 665
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        122 SDKIDE -- SVASYDSSTHERR -- SSRSK -- - KPVNYNSSSE - SESEEQISKATKKVKQKE 173
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     174 EEEYVEEVDEKSLKNESSS-DEFEPVVPEQLETPISKRRSRSSAKNLEKESTMNLDDHA 232
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        718 ----EPVI)ITIKKSDTAGHDSANHVIDASDEKNVSPMGDVPTDTKNEEASVPINRD--A 770
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            233 PREMFDCLDKPIPWRGRLGYACLNTILRSMKERVFCSRTCRITTIQRDGLESVKQLGTQN 292
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     771 TTEVVD-----RELEGETL------RALFEKL------RSELQSLKELTHEK 797
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       293 VLDLIKLVEWNHNFGIHFMRVSSDLFPFASHAKYGYTLEFAQ---SHLEEVGKLANKYNH 349
                                                                                                                                                                                                           'Complete nucleotide sequence of Saccharomyces cerevisiae chromosome
                                                                                                                                                                                                                                                                                                                          Philips J., Herskowitz I.; "Identification of Kellp, a kelch domain-containing protein involved
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                Gaps
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             666 S-KVEEDGNVADEDDEIGVAQMASSPSKDQFKIKHYNESSELSQNNTEIDKLS-----
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            94;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        DB 1; Length 1164;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          64; Mismatches 147; Indels
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               43D0FC570F1D5E4D CRC64;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         KELCH 1.
KELCH 2.
KELCH 3.
KELCH 4.
KELCH 5.
COILED COIL (POTENTIAL).
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           COILED COIL (POTENTIAL)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  _
_
_
                                                                                                                                                                                                                                                                                                                                                                                                                                          -!- SUBUNIT: INTERACTS WITH KEL2.
-!- SIMILARITY: CONTAINS 5 KELCH REPEATS.
-!- SIMILARITY: TO YEAST KEL2.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        4.3%; Score 135;
21.6%; Pred. No. 4
                                  STRAIN=S288C / AB972;
MEDLINE=94378003; PubMed=8091229;
                                                                                                                                                                                                                                                                                                        MEDLINE=99003296; PubMed=9786949;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             1164 AA; 131093 MW;
                                                                                                                                                                                                                                                                                                                                                                                                                          ANTAGONIZE THE PKC1 PATHWAY.
                                                                                                                                                                                                                                             Science 265:2077-2082(1994).
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           EMBL; U10397; AAB68991.1; -.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          PIR; S46769; S46769.
SGD; S0001201; KEL1.
InterPro; IPR001798; Kelch.
Prfam; PF01344; Kelch; 3.
Repeat; Coiled coil.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          84; Conservative
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           307
357
409
460
931
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           1163
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    Best Local Similarity
                SEQUENCE FROM N.A.
                                                                                                                                                                                                                                                                                        CHARACTER/IZATION.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             139
                                                                                                                                                                                          Vaudin M.;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             SEQUENCE
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        Query Match
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           REPEAT
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             REPEAT
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      DOMAIN
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              REPEAT
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    REPEAT
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    REPEAT
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           DOMAIN
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          Matches
Qγ
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     q
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          Öλ
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             qq
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   Öλ
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        qq
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              δ
```

```
This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation-the European Bioinformatics Institute. There are no restrictions on its way non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See http://www.isb-sib.ch/announce/or send an email to license@lsb-sib.ch).
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         28;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            Virology 191:98-105(1992).

-!- PTM: THE N-TERMINAL OF THE COAT PROTEIN IS BLOCKED.

-!- MISCELLANEOUS: THE COAT PROTEIN IS LOCATED IN THE CARBOXY-TERMINAL REGION OF THIS POLYPROTEIN.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          Yoshikawa N., Sasaki E., Kato M., Takahashi T.;
"The nucleotide sequence of apple stem grooving capillovirus genome.";
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               | : : : | : : : | 604 NDCVFKAISAHLGIDŠQDLINFLVNEDISDELLDCIEED---KG-LSHEMIEEVLITKGL 659
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    660 SMVYTSDFKEMAVLNRKYGVNGKMYCTIKGNHCELSSKECFIRLLKEGGEAQMSNENLNA 719
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   -----VNYNSSSESEE 160
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             161 QISKATKKVKQKEEEEYVEEVDEKSLKNESSSDEFEPVVPEQL------ETPISKRRS 213
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        -----IKIEGSAENDQPHEVSHTSMETEDGQG-FEGSIPVDLINCFEPEEIKLPKRRRK 603
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   214 R-----SSAKNLEKESTMN--LDDHAPREMFDCLDKPIPWRGRLGYACLNTIL---- 259
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     291 QNVLDLIKLVEWNHN-----------FGIHF---MRVSSDLFPFASHAKY 326
                                                                   SL------EDKVNELEBIJVNSKFLDIENLNEVI------QFQNEKIKSLELEPNYKEK 893
                                      350 RLTMHPGQYTQIASPREVVVDSAIRDLAYHDEILSRMKLNEQLNKDAVLIIHLGGTFEGK 409
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              Gaps
798 ALEA------GAHIKELETELWQLKSQKNSGTTKEIDELDSVRLQSKCEILEADNH 847
                                                                                                                                                                                                                                                                                                             01-JUN-1994 (Rel. 29, Created)
01-JUN-1994 (Rel. 29, Last sequence update)
01-JUN-2001 (Rel. 40, Last annotation update)
GENOME POLYPROTEIN (CONTAINS: RNA REPLICASE (EC 2.7.7.48); HELICASE;
                                                                                                                                                                                                                                                                                                                                                                                                            Apple stem grooving virus (strain P-209) (ASGV).
Viruses; ssRNA positive-strand viruses, no DNA stage; Capillovirus.
NCBL_TaxID=36402;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      RNA-directed RNA polymerase; Transferase; Polyprotein; ATP-binding;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     ---RVFC-----SRTCRITTIQRDG--LESVKQLGT
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                4.3%; Score 134; DB 1; Length 2105; 21.2%; Pred. No. 10;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            Indels
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              1364 1453 POLYMERASE ACTIVE SITE.
2105 AA; 241240 MW; 331291A5FA137131 CRC64;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            81; Mismatches 193;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             ATP (POTENTIAL)
                                                                                                                                                                                                                                                                      2105 AA
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     119 DSVSDKIDESVASYDSSTHLRRS----SRSKKP---
                                                                                                                         410 KETLDRFRKNYQRLSDSVKARLVLENDDV 438
                                                                                                                                                              894 LEELQIEHENLSRENERLKNESKOHNEDI 922
                                                                                                                                                                                                                                                                      PRT;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      MEDLINE-93033164; PubMed=1413530;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         EMBL; D14995; BAA03639.1; -.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      Best Local Similarity 21.2%
Matches 113; Conservative
                                                                                                                                                                                                                                                                        STANDARD;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         Coat protein; Helicase.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               PIR; A44059; A44059.
MEROPS; C35.001; -.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     SEQUENCE FROM N.A.
                                                                                                                                                                                                                                                                                                                                                                                               PROTEIN].
                                                                                                                                                                                                                                                                        POLR_ASGVP
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        SEQUENCE
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  Query Match
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              551
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       260
                                                                                  848
                                                                                                                                                                                                                                                  POLR_ASGVP
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     ö
                                                                                                                                                                                                                                                                          δ
                                                                                                                                                                g
  a
                                                                             Dp
                                                                                                                         δ
                                           ò
```

```
LLPLCQE-LNI-----PLVLDWHHHNIVPGTLREGSLDLMPLIPTIRETWTRKGITQKQH 498
                                                                                                                                                                                                     327 GYTLEFAQSHLEEVGKLANKYNHRLTMHPGQYTQIASPREVVVDSAIRDLAYHDEILSRM 386
                                                                     777 GLRLGFAGS-----GK-THKVLQWINYTPSVKRMFISPRRMLA-------DEVEPQL 820
                                                                                                       387 KLNE-QLNKDAVLIIHLGGTF-EGKKETLDRFRKNYQRLSDSVKARLVLENDDVSWSVQD 444
720 DSLFDLGRFV---HNRDRAVKLAKSMARGTTGLLNEFDLEFCKNMVTLSELFPENFSSVV 776
                                                                                                                                                                                                                                                                        YSESAD-PTAI--SGMKRRAHSD----RVFDFPPCDPTMDLMIEAKEKEQAVF 544
                                                                                                                                                                                                                                                     499
                                                                                                                                                                                445
                                                                                                                                                                                                                                                                                         636
                                                                                                                                                                            Οÿ
                                                                                                                                                                                                                q
                                                                                                                                                                                                                                                     δ
                                                                 qq
                                                                                                          òλ
                                                                                                                                          g
                                   Qγ
```

Search completed: January 15, 2002, 14:07:05 Job time: 934 sec

. •

.

```
(without alignments)
322.812 Million cell updates/sec
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   Sequence 2, Appli
Sequence 2, Appli
Sequence 8, Appli
Sequence 4, Appli
Sequence 4, Appli
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          App
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             Appl
Appl
Appl
Appl
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                Sequence 10, Appl
Sequence 10, Appl
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         Sequence 2, Appli
Sequence 4, Appli
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             Sequence 6, Appli
Sequence 11, Appl
Sequence 606, App
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   Sequence 11, Appl
Sequence 604, Appl
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          Appli
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  2, Appl
                                                                                                                                                                                                                                                                                                                                                                                                                                                                   Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     603,
                                                                                                                                           4349
1 MTKLPILGYWKIKGLVQPTR......KRLTARKRRSRKEEVEEDEK 828
                                                                           ; Search time 57.72 Seconds
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              Description
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        Sequence 1
Sequence 6
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              Sequence
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  Sequence
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     Sequence
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 Sequence
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               Sequence
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           Sequence
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          Sequence
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    Sequence
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              Sequence
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  Sequence
                                                                                                                                                                                                                                                                                                                                                                            /cgn2_6/ptodata/2/iaa/5A_COMB.pep:*
/cgn2_6/ptodata/2/iaa/5B_COMB.pep:*
/cgn2_6/ptodata/2/iaa/6B_COMB.pep:*
/cgn2_6/ptodata/2/iaa/6B_COMB.pep:*
/cgn2_6/ptodata/2/iaa/PCTUS_COMB.pep:*
/cgn2_6/ptodata/2/iaa/PCTUS_COMB.pep:*
GenCore version 4.5
Copyright (c) 1993 - 2000 Compugen Ltd
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      US-08-395-507-1
US-08-974-549A-603
US-08-129-722A-2
PCT-US94-09700-6
PCT-US94-09700-11
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               US-08-974-549A-604
US-08-974-549A-605
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      US-08-974-549A-606
US-08-974-549A-602
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               US-09-352-159-19
US-09-352-168-19
US-09-352-168-33
US-09-352-168-33
US-09-352-159-31
US-09-352-159-29
US-09-352-159-29
                                                                                                                                                                                                                                               Total number of hits satisfying chosen parameters:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             US-08-216-894-8
US-09-115-746-8
US-08-216-894-4
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    US-09-115-746-4
US-08-216-894-10
US-09-115-746-10
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         US-08-395-507-2
US-08-737-248-4
US-08-864-224-11
                                                                                                                                                                                                                     212252 seqs, 22503292 residues
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         SUMMARIES
                                                                         January 15, 2002, 13:45:31
                                                                                                                                                                                                                                                                                                             Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries
                                                    protein search, using sw model
                                                                                                                                                                                                Gapop 10.0 , Gapext 0.5
                                                                                                                                                                                                                                                                                                                                                                   Issued_Patents_AA:*
                                                                                                                                                                                                                                                                          Minimum DB seq length: 0
Maximum DB seq length: 200000000
                                                                                                                               US-09-724-296-2
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               DB
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               Query
Match Length D
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    232
232
332
332
332
332
                                                                                                                                                                                  BLOSUM62
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       1200.5
11999.5
11999.5
11999.5
11999.5
11999.5
11999.5
11999.5
11999.5
11999.5
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               Score
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      1259.5
1259.5
1259.5
1259.5
1228.5
1228.5
1221.5
1221.5
1207.5
1206.5
1205.5
                                                                                                                                           Perfect score:
                                                                                                                                                                                  Scoring table:
                                                     OM protein
                                                                                                                                                                                                                         Searched:
                                                                                                                                                          Sequence:
                                                                                                                                                                                                                                                                                                                                                                    Database
                                                                              Run on:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              Result
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               ٠
9
```

```
Sequence 1, Appli
Sequence 56, Appl
Sequence 6, Appli
Sequence 9, Appli
Sequence 20, Appli
Sequence 4, Appli
Sequence 6, Appli
                                                                                                                         Appli
Appli
Appli
                                                                                                                                                                         , Appl
Appli
                                                                                                                                                                                                      Sequence 26, Appl
Sequence 33, Appl
Patent No. 5427917
                                                                                                                                                                                                                                                     Sequence 4, Appli
Sequence 6, Appli
                                                                                                                                       Sequence 3, Ap
Sequence 5, Ap
Sequence 11, A
Sequence 2, Ap
Sequence 29,
                                                                                                                                                                                                                                                                                                                                                                              US-08-216-894-2
Sequence 2, Application US/08216894
Sequence 2, Application US/08216894
Patent No. 5876734
GENERAL INFORMATION:
APPLICANT: Rirchhoff, Louis V.
APPLICANT: Otsu, Keiko
TITLE OF INVENTION: POLYPEPTIDES FOR DIAGNOSING INFECTION
TITLE OF INVENTION: WITH TRYPANOSOMA CRUZI
NUMBER OF SEQUENCES: 10
CORRESPONDENCE ADDRESS:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        SOFTWARE: Patentin Release #1.0, Version #1.25 CURRENT APPLICATION DATA:
         PCT - US95 - 04439 - 1

US - 08 - 216 - 894 - 6

US - 09 - 115 - 746 - 6

US - 09 - 115 - 746 - 6

US - 08 - 102 - 757 - 9

US - 08 - 910 - 820 - 6

US - 08 - 910 - 820 - 6

US - 08 - 910 - 820 - 6

US - 08 - 910 - 820 - 5

US - 08 - 910 - 820 - 5

US - 08 - 910 - 820 - 5

US - 08 - 910 - 820 - 5

US - 08 - 910 - 820 - 5

US - 08 - 910 - 820 - 5

US - 08 - 910 - 820 - 5

US - 08 - 910 - 820 - 5

US - 08 - 910 - 820 - 5

US - 08 - 910 - 820 - 5

US - 08 - 910 - 820 - 5

US - 08 - 910 - 820 - 5
                                                                                                                                                                                                                       US-08-297-431B-33
                                                                                                                                                                                                        US-08-470-837-26
                                                                                                                                                                                                                                                       US-08-747-536-4
US-08-747-536-6
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    REFERENCE/DOCKET NUMBER: 85326/102/DRLO TELECOMMUNICATION:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          ADDRESSEE: Foley & Lardner
STREET: 3000 K Street, N.W., Suite 500
CITY: Washington, D.C.
COUNTRY: USA
                                                                                                                                                                                                                                                                                                                       ALIGNMENTS
                                                                                                                                                                                                                                         5427917-2
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     APPLICATION NUMBER: US/08/216,894
FILING DATE: 24-MAR-1994
ATTORNEY/AGENT INFORMATION:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      NAME: BENT, Stephen A. REGISTRATION NUMBER: 29,768
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  TELEPHONE: (202)672-5300
TELERAX: (202)672-5399
INFORMATION FOR SEQ ID NO: 2:
SEQUENCE CHARACTERISTICS:
LENGTH: 564 amino acids
TYPE: amino acid
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            Floppy disk
              435
647
307
307
307
2291
272
272
272
272
272
272
272
270
210
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         ZIP: 20007-5109
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy C
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              ; MOLECULE TYPE: protein US-08-216-894-2
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   linear
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 TOPOLOGY:
                                                                                                        1180.5
1180.5
1180.5
1180.5
1179
1172.5
1170
471.5
                                              1194
1194
1193
1187
                                                                                                                                                                                                                                        241
240
240
1199.5
 228
332
332
332
332
444
447
447
447
447
447
447
```

12;

5 PILGYWKIKGLVQPTRLLLEYLEEKYEEHLYERDEGDKWRNKKFELGLEFPNLPYYIDGD 64

29.0%; Score 1259.5; DB 2; Length 564; 57.1%; Pred. No. 1.9e-94; Live 55; Mismatches 103; Indels 45;

Best Local Similarity 57.13 Matches 270; Conservative

Oy Db

Query Match

```
301 INLKKRPNDAVSNRDKKKNSETAKTD-EVEKQRAAEAAKAVETEKQRAAEAT-KVAEAEK 358
185 IEAIPQIDKYLKSSKYIAWPLQGWQATFGGGDHPPKSDHLVPRG--SMLRLLKRNIQISK 242
                                                                                                                                                                    242 K-EMTKLREKVKKAEKEKLDAINRATKLEEERNQAYKAAHKAEEEKAKTFQRLITFESEN 300
                                                                                                                                                                                                                                                                                                             290 -TLSMLPQVNIG-----ANSFSAETPVDLKKENETELANISGPHKKSTSTSTRKRARSSK 343
                                                                                                                                                                                                                                                                                                                                                                                            344 KKATDSV----SDKIDESVASYDSSTHLRRSSRSKKPVNYNSSSESESEEQISKATKKVK 399
                                                                                                                                                                                                                                                                                                                                                                                                                                     359 RKAAEAAKAVETEKQRAAEATKVAEAEKQKAAEAAKAV------ETEKQRAAEATKVA 410
                                                               125 LSKLPEMLKMFEDRLCHKTYLNGDHVTHPDFMLYDALDVVLYMDPMCLDAFPKLVCFKKR 184
                                                                                                                                                                                                                             243 RIVFTILKQKAFKGNHPCVPSVCTIT-----YSRFHCLPD----TLKSLLPMSSKT 289
                                                                                                                                                                                                                                                                                                                                                                                                                                                                             400 QKEEEEYVEEVDEKSLKNESSSDEFEPVVPEQLETPISKRRSRSSAKNLEKE 452
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     411 EAEKORAAEAMKVAEAEKOKAA-----EATKVAEAEKOKAAEATKVAEAE 455
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            TITLE OF INVENTION: POLYPEPTIDES FOR DIAGNOSING INFECTION TITLE OF INVENTION: WITH TRYPANOSOMA CRUZI NUMBER OF SEQUENCES: 10
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentin Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 ATTORNEY/AGENT INFORMATION:
NAME: BENT, Stephen A.
REGISTRATION NUMBER: 29,768
REFERENCE/DOCKET NUMBER: 85326/102/DRLO
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                SEE: Foley & Lardner: 3000 K Street, N.W., Suite 500 Washington, D.C.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  APPLICATION NUMBER: US/09/115,746
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      PRIOR APPLICATION DATA: APPLICATION NUMBER: 08/216,894
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            Sequence 2, Application US/09115746 Patent No. 6228601
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     APPLICANT: Kirchhoff, Louis V. APPLICANT: Otsu, Keiko
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           E: Floppy disk
IBM PC compatible
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            TELERAX: (202)672-5399
TELEX: 904136
INFORMATION FOR SEQ ID NO: 2:
SEQUENCE CHARACTERISTICS:
LENGTH: 564 amino acids
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             (202)672-5300
(202)672-5399
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            CORRESPONDENCE ADDRESS:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     COMPUTER READABLE FORM: MEDIUM TYPE: Floppy
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         MOLECULE TYPE: protein
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     amino acid
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      GENERAL INFORMATION:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              linear
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               20007-5109
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  USA
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    FILING DATE:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      FILING DATE
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    ADDRESSEE:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          TELEPHONE:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 COMPUTER:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         TOPOLOGY:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    US-09-115-746-2
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               US-09-115-746-2
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                COUNTRY:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        STREET:
                                                                                                     q
                                                                                                                                                                                    Ω
                                                                                                                                                                                                                                                                     qq
                                                                                                                                                                                                                                                                                                                                                    qq
                                                                                                                                                                                                                                                                                                                                                                                                                                 g
                                                               ŏ
                                                                                                                                                                                                                                ŏ
                                                                                                                                                                                                                                                                                                             ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                             οy
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  g
                                                                                                                                                                                                                                                                                                                                                                                              ò
```

```
12;
                                                                                                                                                                                  63 VKLTQSMAITRYIADKHNMLGGCPKERAEISMLEGAVLDIRYGVSRIAYSKDFETLKVDF 122
                                                                                                                                                                                                                                                                        123 LSKLPEMLKMFEDRLCHKTYLNGDHVTHPDFMLYDALDVVLYMDPMCLDAFPKLVCFKKR 182
                                                                                                                                                                                                                                                                                                                                                                242 K-EMTKLEEKVKKAEKEKLDAINRATKLEEERNQAYKAAHKAEEEKAKTFORLITFESEN 300
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   45; Gaps
                                                                                                                                                           65 VKLTQSMALIRYIADKHNMLGGCPKERAEISMLEGAVLDIRYGVSRIAYSKDFETLKVDF 124
                                                                                                                                                                                                                                                   125 LSKLPEMLKMFEDRLCHKTYLNGDHVTHPDFMLYDALDVVLYMDPMCLDAFPKLVCFKKR 184
                                                                                                                                                                                                                                                                                                                                         185 IEAIPQIDKYLKSSKYIAWPLQGWQATFGGGDHPPKSDHLVPRG--SMLRLLKRNIQISK 242
                                                                                                                                                                                                                                                                                                                                                                                                                              243 RIVFTILKQKAFKGNHPCVPSVCTIT-----YSRFHCLPD----TLKSLLPMSSKT 289
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       290 -TLSMLPQVNIG-----ANSFSAETPVDLKKENETELANISGPHKKSTSTSTRKRARSSK 343
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               344 KKATDSV----SDKIDESVASYDSSTHLRRSSRSKKPVNYNSSSESEEGISKATKKVK 399
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       359 RKAAEAAKAVETEKQRAAEATKVAEAEKOKAAEAAKAV-----ETEKORAAEATKVA 410
                                                                                          3 PIGGYWKIKGLVQPTRLLLEYLEEKYEEHLYERDGGDKWRNKKFELGLEFPULPYYIDGD 62
                                                                  5 PILGYWKIKGLVQPTRLLLEYLEEKYEEHLYERDEGDKWRNKKFELGLEFPNLPYYIDGD 64
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       400 QKEEEEYVEEVDEKSLKNESSSDEFEPVVPEQLETPISKRRSRSSAKNLEKE 452
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                411 EAEKQRAAEAMKVAEAEKQKAA-----EATKVAEAEKQKAAEATKVAEAE 455
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              POLYPEPTIDES FOR DIAGNOSING INFECTION WITH TRYPANOSOMA CRUZI
57.1%; Pred. No. 1.9e-94; ative 55; Mismatches 103; Indels
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       SOFTWARE: Patentin Release #1.0, Version #1.25 CURRENT APPLICATION DATA:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         85326/102/DRLO
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      SEE: Foley & Lardner: 3000 K Street, N.W., Suite 500 Washington, D.C.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         US/08/216,894
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         PC-DOS/MS-DOS
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                Sequence 8, Application US/08216894 Patent No. 5876734
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              APPLICANT: Kirchhoff, Louis V. APPLICANT: Otsu, Keiko
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               29,768
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          TELECOMMUNICATION INFORMATION:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          TELEX: 904136
INFORMATION FOR SEQ ID NO: 8;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          FILING DATE: 24-MAR-1994
ATTORNEY/AGENT INFORMATION:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  (202)672-5300
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    REFERENCE/DOCKET NUMBER:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        SEQUENCE CHARACTERISTICS:
LENGTH: 643 amino acids
                            Conservative
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       (202)672-5399
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         NAME: BENT, Stephen A. REGISTRATION NUMBER: 2
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           10
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            CORRESPONDENCE ADDRESS:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         APPLICATION NUMBER:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           NUMBER OF SEQUENCES:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 COMPUTER: IBM PC
OPERATING SYSTEM:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           TITLE OF INVENTION:
  Best Local Similarity
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     amino acid
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         20007-5109
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           GENERAL INFORMATION:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    TELEPHONE:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      ADDRESSEE:
                          Matches 270;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          US-08-216-894-8
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         COUNTRY:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              STREET:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       RESULT
                                                                                                             qq
                                                                                                                                                             ò
                                                                                                                                                                                                     qq
                                                                                                                                                                                                                                                   δ
                                                                                                                                                                                                                                                                                              g
                                                                                                                                                                                                                                                                                                                                                                                    q
                                                                                                                                                                                                                                                                                                                                                                                                                                 ōλ
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       δ
                                                                                                                                                                                                                                                                                                                                                                                                                                                                            g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          qq
```

linear

TOPOLOGY:

29.0%; Score 1259.5; DB 4; Length 564;

Query Match

```
185
                                                                                                                                                                                                                                                                                                                                                                                                                            125
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    RESULT
                                                                                                                                                                                                                                                                                                                      qq
                                                                                                                                                                                                                                                                                      õ
                                                                                                                                                                                                                                                                                                                                                           δ
                                                                                                                                                                                                                                                                                                                                                                                          pp
                                                                                                                                                                                                                                                                                                                                                                                                                            Óλ
                                                                                                                                                                                                                                                                                                                                                                                                                                                               ρp
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  q
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     οy
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     q_0
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         Pρ
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              Ο
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             QQ
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 Qγ
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          δ
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            q
                                                                                                           12;
                                                                                                                                                                                                                                                                                                                                                                183 IEAIPQIDKYLKSSKYIAWPLQGWQATFGGGDHPPKSD-LVPRGSPSQLQQAENNITNSK 241
                                                                                                                                                                                                                                 123 LSKLPEMLKNFEDRLCHKTYLNGDHVTHPDFMLYDALDVVLYMDPMCLDAFPKLVCFKKR 182
                                                                                                                                                                                                                                                                                                                                                       IEAIPQIDKYLKSSKYIAWPLQGWQATFGGGDHPPKSDHLVPRG--SMLRLLKRNIQISK 242
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     65 VKLTQSMAIIRYIADKHNMLGGCPKERAEISMLEGAVLDIRYGVSRIAYSKDFETLKVDF 124
                                                                                                                                                                                                                                                                                  LSKLPEMLKMFEDRLCHKTYLNGDHVTHPDFMLYDALDVVLYMDPMCLDAFPKLVCFKKR 184
                                                                                                                                                                                                                                                                                                                                                                                                                         243 RIVFTILKQKAFKGNHPCVPSVCTIT-----YSRFHCLPD----TLKSLLPMSSKT 289
                                                                                                                                                                                                                                                                                                                                                                                                                                                         242 K-EMTKLREKVKKAEKEKLDAINRATKLEEERNQAYKAAHKAEEEKAKTFORLITFESEN 300
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          290 -TLSMLPQVNIG-----ANSFSAETPVDLKKENETELANISGPHKKSTSTSTRKRARSSK 343
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               344 KKATDSV----SDKIDESVASYDSSTHLRRSSRSKKPVNYNSSSESEEGISKATKKVK 399
                                                                                                                                                            3 PILGYWKIKGLVQPTRLLLEYLEEKYEEHLYERDEGDKWRNKKFELGLEFPULPYYIDGD 62
                                                                                                             Gaps
                                                                                                                                          5 PILGYWKIKGLVQPTRLLLEYLEEKYEEHLYERDEGDKWRNKKFELGLEFPNLPYYIDGD 64
                                                                                                           45;
                                                                          Length 643;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  400 QKBEEEYVEEVDEKSLKNESSSDEFEPVVPEQLETPISKRRRSRSSAKNLEKE 452
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   411 EAEKORAAEAMKVAEAEKOKAA-----EATKVAEAEKOKAAEATKVAEAE 455
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                DIAGNOSING INFECTION
                                                                                                       55; Mismatches 103; Indels
                                                                     29.0%; Score 1259.5; DB 2; 57.1%; Pred. No. 2.3e-94;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        PatentIn Release #1.0, Version #1.25
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          APPLICANT: Kirchhoff, Louis V.
APPLICANT: Otsu, Keiko
TITLE OF INVENTION: POLYPEPTIDES FOR DIAGNO
TITLE OF INVENTION: WITH TRYPANOSOMA CRUZI
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 REFERENCE/DOCKET NUMBER: 85326/102/DRLO
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            3000 K Street, N.W., Suite 500
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            APPLICATION NUMBER: US/09/115,74
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           PC-DOS/MS-DOS
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             PRIOR APPLICATION DATA:
APPLICATION NUMBER: 08/216,894
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        Sequence 8, Application US/09115746 Patent No. 6228601
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  COMPUTER: IBM PC compatible OPERATING SYSTEM: PC-DOS/MS-
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                NAME: BENT, Stephen A. REGISTRATION NUMBER: 29,768
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   TELECOMMUNICATION INFORMATION:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 ADDRESSEE: Foley & Lardner
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  ATTORNEY/AGENT INFORMATION:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       Floppy disk
                                                                                                       Matches 270; Conservative
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  10
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  Washington, D.C.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       CURRENT APPLICATION DATA
MOLECULE TYPE: protein
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           NUMBER OF SEQUENCES: 1/CORRESPONDENCE ADDRESS:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     COMPUTER READABLE FORM: MEDIUM TYPE: Floppy
                                                                                         Similarity
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       GENERAL INFORMATION:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    20007-5109
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      USA
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               FILING DATE:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  FILING DATE:
     ; MOLECULE TY
US-08-216-894-8
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          RESULT 4
US-09-115-746-8
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    COUNTRY:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   STREET:
                                                                      Query Match
                                                                                       Best Local
                                                                                                                                                                                                                                                                                125
                                                                                                                                                                                                                                                                                                                                                   185
                                                                                                                                                                                                                                              g
                                                                                                                                                                                                                                                                                                                 g
                                                                                                                                                                                                                                                                                                                                                                                  g
                                                                                                                                                                                                                                                                                                                                                                                                                                                       Q
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             g
                                                                                                                                          ŏ
                                                                                                                                                                          g
                                                                                                                                                                                                             δ
                                                                                                                                                                                                                                                                                ò
                                                                                                                                                                                                                                                                                                                                                                                                                       ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            δ
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               õ
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 g
                                                                                                                                                                                                                                                                                                                                                     ò
```

```
12;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    TEAIPQIDKYLKSSKYIAWPLQGWQATFGGGDHPPKSDHLVPRG--SMLRLLKRNIQISK 242
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         243 RIVFTILKQKAFKGNHPCVPSVCTIT-----YSRFHCLPD----TLKSLLPMSSKT 289
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  242 K-EMTKLREKVKKAEKEKLDAINRATKLEEERNQAYKAAHKAEEEKAKTPQRLITFESEN 300
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     65 VKLTQSMAIIRYIADKHNMLGGCPKERAEISMLEGAVLDIRYGVSRIAYSKDFETLKVDF 124
                                                                                                                                                                                                                                                                                                                                                                                                                                                               LSKLPEMLKMFEDRLCHKTYLNGDHVTHPDFMLYDALDVVLYMDPMCLDAFPKLVCFKKR 184
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            290 -TLSMLPQVNIG-----ANSFSAETPVDLKKENETELANISGPHKKSTSTSTRKRARSSK 343
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 344 KKATDSV----SDKIDESVASYDSSTHLRRSSRSKKPVNYNSSSESEEQISKATKKVK 399
                                                                                                                                                                                                                                                                                                            Gaps
                                                                                                                                                                                                                                                                                                                                                                        5 PILGYWKIKGLVQPTRILLEYLEEKYEEHLYERDEGDKWRNKKFELGLEFPNLPYYIDGD 64
                                                                                                                                                                                                                                                                                                        45;
                                                                                                                                                                                                                                                              DB 4; Length 643;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     400 QKEEEEYVEEVDEKSLKNESSSDEFEPVVPEQLETPISKRRRSRSSAKNLEKE 452
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            411 EAEKQRAAEAMKVAEAEKQKAA-----EATKVAEAEKOKAAEATKVAEAE 455
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    TITLE OF INVENTION: POLYPEPTIDES FOR DIAGNOSING INFECTION TITLE OF INVENTION: WITH TRYPANOSOMA CRUZI
                                                                                                                                                                                                                                                                                                        103; Indels
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          SOFTWARE: PatentIn Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/216,894
FILING DATE: 24-MAR-1994
                                                                                                                                                                                                                                                            Score 1259.5; DB Pred. No. 2.3e-94;
                                                                                                                                                                                                                                                                                                          Mismatches
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              E: Foley & Lardner
3000 K Street, N.W., Suite 500
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        Sequence 4, Application US/08216894 Patent No. 5876734 GENERAL INFORMATION:
                                                                                                                                                                                                                                                                                                     55;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         APPLICANT: Kirchhoff, Louis V. APPLICANT: Otsu, Keiko
                                                                                                                                                                                                                                                            29.0%;
             TELEFAX: (202)672-5399
TELEX: 904136
INFORMATION FOR SEQ ID NO: 8:
(202)672-5300
                                                                              SEQUENCE CHARACTERISTICS:
LENGTH: 643 amino acids
TYPE: amino acid
                                                                                                                                                                                                                                                                                                        Matches 270; Conservative
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  10
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         Washington, D.C.
                                                                                                                                                                     ; MOLECULE TYPE: protein US-09-115-746-8
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             NUMBER OF SEQUENCES: 10
CORRESPONDENCE ADDRESS:
ADDRESSEE: Foley & La
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                ZIP: 20007-5109
COMPUTER READABLE FORM:
                                                                                                                                                     linear
                                                                                                                                                                                                                                                       Query Match
Best Local Similarity
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              USA
relephone:
                                                                                                                                                  TOPOLOGY:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    US-08-216-894-4
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              COUNTRY:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                STREET:
```

```
RESULT.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           q
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   113 LSKLPEMLKMFEDRLCHKTYLNGDHVTHPDFMLYDALDVVLYMDPMCLDAFPKLVCFKKR 182
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          243 RIVFTILKQKAFKGNHPCVPSVCTIT-----YSRFHCLPD----TLKSLLPMSSKT 289
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             290 TLSMLPQVNIGANSFSAETPVDLKKENETELANISGPHKKSTSTSTRKRARSSKKKATDS 349
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         300 -----NINLKKRPNDAVSNRDKKKNSETAKTD---EVEKQRAAEAAKAVETEKQRAGEF 350
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    63 VKLTQSMAIIRYIADKHNMLGGCPKERAEISMLEGAVLDIRYGVSRIAYSKDFETLKVDF 122
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  12 LSKLPEMLKMFEDRLCHKTYLNGDHVTHPDFMLYDALDVVLYMDPMCLDAFPKLVCFKKR 184
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    242 K-EMTKLREKVKKAEKEKLDAINRATKLEEERNQAYKAAHKAEEEKAKTFQRLITFESE- 299
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             65 VKLTQSMAIIRYIADKHNMLGGCPKERAEISMLEGAVLDIRYGVSRIAYSKDFETLKVDF 124
                                                                                                                                                                                                                                                                                                                                                                                                              27; Gaps
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 3 PILGYWKIKGLVQPTRLLLEYLEEKYEEHLYERDEGDKWRNKKFELGLEFPNLPYYLDGD 62
                                                                                                                                                                                                                                                                                                                                                                                                                                                       5 PILGYWKIKGLVQPTRLLLEYLEEKYEEHLYERDEGDKWRNKKFELGLEFPNLPYYIDGD 64
                                                                                                                                                                                                                                                                                                                                                                     DB 2; Length 354;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             TITLE OF INVENTION: POLYPEPTIDES FOR DIAGNOSING INFECTION TITLE OF INVENTION: WITH TRYPANOSOMA CRUZI NUMBER OF SEQUENCES: 10
                                                                                                                                                                                                                                                                                                                                                                                                              60; Indels
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           SOFTWARE: Patentin Release #1.0, Version #1.25 CURRENT APPLICATION DATA:
                                                                                                                                                                                                                                                                                                                                                                  28.2%; Score 1228.5; DB 369.3%; Pred. No. 3.1e-92; iive 24; Mismatches 60;
                NAME: BENT, Stephen A. REGISTRATION NUMBER: 29,768
REGISTRATION NUMBER: 85326/102/DRLO
TELECOMMUNICATION INFORMATION:
TELEPHONE: (202)672-5300
TELEFAX: (202)672-5399
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   ADDRESSEE: Foley & Lardner
STREET: 3000 K Street, N.W., Suite 500
CITY: Washington, D.C.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         APPLICATION NUMBER: US/09/115,746
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               Sequence 4, Application US/09115746 Patent No. 6228601
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  APPLICANT: Kirchhoff, Louis V. APPLICANT: Otsu, Keiko
                                                                                                              TELEX: 904136
INFORMATION FOR SEQ ID NO: 4:
SEQUENCE CHARACTERISTICS:
TENCH: 354 amir-
ATTORNEY/AGENT INFORMATION:
                                                                                                                                                                                                                                                                                                                                                                                                              Matches 250; Conservative
                                                                                                                                                                                                                                                                               MOLECULE TYPE: protein
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              CORRESPONDENCE ADDRESS:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          COMPUTER READABLE FORM:
                                                                                                                                                                                                                                                                                                                                                                     Query Match
Best Local Similarity
                                                                                                                                                                                                                                                             linear
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               GENERAL INFORMATION:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 20007-5109
                                                                                                                                                                                                                                                        TOPOLOGY:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  350 V 350
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 US-09-115-746-4
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            351 I 351
                                                                                                                                                                                                                                                                                                    US-08-216-894-4
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 COUNTRY:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              RESULT
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  δŏ
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 Db
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               ŏ
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                δλ
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            QQ
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          ŏ
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    q
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             Qγ
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    Dp
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         g
```

```
63 VKLTQSMAIIRYIADKHNMLGGCPKERAEISMLEGAVLDIRYGVSRIAYSKDFETLKVDF 122
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         123 LSKLPEMLKMFEDRLCHKTYLNGDHVTHPDFMLYDALDVVLYMDPMCLDAFPKLVCFKKR 182
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              185 IEAIPQIDKYLKSSKYIAWPLQGWQATFGGGDHPPKSDHLVPRG--SMLRLLKRNIQISK 242
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     243 RIVFTILKQKAFKGNHPCVPSVCTIT-----YSRFHCLPD----TLKSLLPMSSKT 289
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      242 K-EMTKLREKVKKAEKEKLIJAINRATKLEEERNQAYKAAHKAEEEKAKTFORLITFESE- 299
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      290 TLSMLPQVNIGANSFSAETPVDLKKENETELANISGPHKKSTSTSTRRKRARSSKKKATDS 349
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  300 -----NINLKKRPNDAVSNRDKKKNSETAKTD----EVEKQRAAEAAKAVETEKQRAGEF 350
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 65 VKLTQSMAIIRYIADKHNMLGGCPKERAEISMLEGAVLDIRYGVSRIAYSKDFETLKVDF 124
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  LSKLPEMLKMFEDRLCHKTYLNGDHVTHPDFMLYDALDVVLYMDPMCLDAFPKLVCFKKR 184
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        Gaps
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            3 PILGYWKIKGLVQPTRLLLEYLEEKYEEHLYERDEGDKWRNKKFELGLEFPNLPYYIDGD 62
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     5 PILGYWKIKGLVOPTRLLLEYLEKYEEHLYERDEGDKWRNKKFELGLEFPNLPYYIDGD 64
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    28.2%; Score 1228.5; DB 4; Length 354; 69.3%; Pred. No. 3.1e-92;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           TITLE OF INVENTION: POLYPEPTIDES FOR DIAGNOSING INFECTION TITLE OF INVENTION: WITH TRYPANOSOMA CRUZI
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        60; Indels
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     24; Mismatches
                                                                                                                                                              85326/102/DRLO
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                3000 K Street, N.W., Suite 500
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   Sequence 10, Application US/08216894; Patent No. 5876734; GENERAL INFORMATION:
               PRIOR APPLICATION DATA: APPLICATION NUMBER: 08/216,894
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             APPLICANT: Kirchhoff, Louis V. APPLICANT: Otsu, Keiko
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          IBM PC compatible
                                                                                                                                       29,768
                                                                                                                                                           REFERENCE/DOCKET NUMBER: 85
TELECOMMUNICATION INFORMATION:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            Foley & Lardner
                                                                                                                                                                                                                                   TELEFAX: (202)672-5399
TELEX: 904136
INFORMATION FOR SEQ ID NO: 4:
                                                                                      ATTORNEY/AGENT INFORMATION:
                                                                                                                                                                                                              (202)672-5300
(202)672-5399
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   Floppy disk
                                                                                                                                                                                                                                                                                                                                    354 amino acids
                                                                                                                                                                                                                                                                                                     SEQUENCE CHARACTERISTICS:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        Matches 250; Conservative
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       Washington, D.C.
                                                                                                                  NAME: BENT, Stephen A. REGISTRATION NUMBER: 2
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              10
                                                                                                                                                                                                                                                                                                                                                                                                   ; MOLECULE TYPE: protein US-09-115-746-4
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            COMPUTER READABLE FORM: MEDIUM TYPE: Floppy
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  CORRESPONDENCE ADDRESS:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              NUMBER OF SEQUENCES:
                                                                                                                                                                                                                                                                                                                                                        amino acid
                                                                                                                                                                                                                                                                                                                                                                                       linear
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 Best Local Similarity
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     20007-5109
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   USA
FILING DATE:
                                                                  FILING DATE:
                                                                                                                                                                                                              TELEPHONE:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            ADDRESSEE:
                                                                                                                                                                                                                                                                                                                                                                             TOPOLOGY:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               COMPUTER:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              US-08-216-894-10
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 COUNTRY:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       350 V 350
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  351 I 351
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   STREET:
                                                                                                                                                                                                                                                                                                                                    LENGTH:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       Query Match
```

ഹ

Page

```
RESULT
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             Op
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        q
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       QQ
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    a
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             ŏ
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       δ
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             q
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          6
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      342 KAAEATKVAEAEKOKAAEATKVAEAEKOKAAEATKVAEAEKOKAAEATKVAEAEKOKAA - 400
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 63 VKLTQSMAIIRYIADKHNMLGGCPKERAEISMLEGAVLDIRYGVSRIAYSKDFETLKVDF 122
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    363 SSTHLRRSSRSKKPVNYNSSSESESEEQISKATKKVKQKEEEEFYVEEVDEKSLKNESSSD 422
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    65 VKLTQSMAIIRYIADKHNMLGGCPKERAEISMLEGAVLDIRYGVSRIAYSKDFETLKVDF 124
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              LSKLPEMLKMFEDRLCHKTYLNGDHVTHPDFMLYDALDVVLYMDPMCLDAFPKLVCFKKR 184
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             185 IEAIPQIDKYLKSSKYIAWPLQGWQATFGGGDHPPKSDHLVPRGSMLRLLKRNIQISKRI 244
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       245 VFTILKQKAFKGNHPCVPSVCTITYSRFHCLPDTLKSLLPMSSKTTLSMLPQVNIGANSF 304
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      238 A-EAEKQKAAEAT-----KVAEAEKQRAAEATK--VAEAEKQKAAEATKVAEAEKQR 286
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 SAETP--VDLKKENETELANISGPHKKSTSTSTRKRARSSKKKATDSVSDKIDESVASYD 362
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       3 PILGYWKIKGLVQPTRLLLEYLEEKYEEHLYERDEGDKWRNKKFELGLEFPNLPYYIDGD 62
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          31; Gaps
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     5 PILGYWKIKGLVQPTRLLLEYLEEKYEEHLYERDEGDKWRNKKFELGLEFPNLPYYIDGD 64
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           Query Match 28.1%; Score 1221.5; DB 2; Length 472; Best Local Similarity 56.7%; Pred. No. 1.8e-91; Matches 255; Conservative 54; Mismatches 110; Indels 31;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              TITLE OF INVENTION: POLYPEPTIDES FOR DIAGNOSING INFECTION TITLE OF INVENTION: WITH TRYPANOSOMA CRUZI NUMBER OF SEQUENCES: 10 CORRESPONDENCE ADDRESS:
                    SOFTWARE PATENTIAN STATES OF THE STATES OF T
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  3000 K Street, N.W., Suite 500
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          423 EFEPVVPEQLETPISKRRRSRSSAKNLEKE 452
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      401 -----EATKVAEAEKOKAAEATKVAEAE 423
PC-DOS/MS-DOS
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 Sequence 10, Application US/09115746
Patent No. 6228601
GENERAL INFORMATION:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   APPLICANT: Kirchhoff, Louis V. APPLICANT: Otsu, Keiko
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 Foley & Lardner
                                                                                                                                                                                                                                                                                                                                                                                                                                                                  TELEX: 904136
INFORMATION FOR SEQ ID NO: 10:
SEQUENCE CHARACTERISTICS:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               I: 472 amino acids
amino acid
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                CITY: Washington, D.C.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      ; TOPOLOGY: linear
; MOLECULE TYPE: protein
US-08-216-894-10
OPERATING SYSTEM:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            ADDRESSEE:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    US-09-115-746-10
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               STREET:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               LENGTH:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  125
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       οy
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 õ
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             a
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               QQ
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  ò
```

```
305 SAETP--VDLKKENETELANISCPHKKSTSTRTRRARSSKKKATDSVSDKIDESVASYD 362
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          363 SSTHLRRSSRSKKPVNYNSSSESEEGISKATKKVKQKEEEEFVEEVDEKSLKNESSSD 422
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               342 KAAEATKVAEAEKQKAAEATKVAEAEKQKAAEATKVAEAEKQKAAEATKVAEAEKQKAA- 400
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        183 IEAIPQIDKYLKSSKYIAWPLQGWQATFGGGDHPPKSDLIBGRG-----IPPGCRNSTKV 237
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         65 VKLTQSMAIIRYIADKHNMLGGCPKERAEISMLEGAVLDIRYGVSRIAYSKDFETLKVDF 124
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               63 VKLTQSMAIIRYIADKHNMLGGCPKERAEISMLEGAVLDIRYGVSRIAYSKDFETLKVDF 122
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            LSKLPEMLKMFEDRLCHKTYLNGDHVTHPDFMLYDALDVVLYMDPMCLDAFPKLVCFKKR 184
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     185 IEAIPQIDKYLKSSKYIAWPLQGWQATFGGGDHPPKSDHLVPRGSMLRLLKRNIQISKRI 244
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              245 VFTILKQKAFKGNHPCVPSVCT1TYSRFHCLPDTLKSLLPMSSKTTLSMLPQVNIGANSF 304
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           3 PILGYWKIKGLVQPTRLLLEYLEEXYEEHLYERDEGDKWRNKKFELGLEFPNLPYYIDGD 62
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                5 PILGYWKIKGLVQPTRLLLEYLEEKYEEHLYERDEGDKWRNKKFELGLEFPNLPYYIDGD 64
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      287 AAEATKVAEAEKQKAAEATKVAGDEKQKAAEAT-KVAEAEKQKAAEAT--KVAE--AEKQ
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             238 A-EAEKQKAAEAT-----KVAEAEKQRAAEATK--VAEAEKQKAAEATKVAEAEKQR
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            DB 4; Length 472;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              Indels
                                                                                                                                                     PatentIn Release #1.0, Version #1.25
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         110;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       28.1%; Score 1221.5; DB 56.7%; Pred. No. 1.8e-91 Live 54; Mismatches 11
                                                                                                                                                                                                                                                                                                                                                                                                                   85326/102/DRLO
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      423 EFEPVVPEQLETPISKRRRSRSSAKNLEKE 452
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        -----EATKVAEAEKQKAAEATKVAEAE 423
                                                                                                                                                                           CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/09/115,746
                                                                                             COMPUTER: IBM PC compatible OPERATING SYSTEM: PC-DOS/MS-DOS
                                                                                                                                                                                                                                                  PRIOR APPLICATION DATA:
APPLICATION NUMBER: 08/216,894
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    ; Sequence 2, Application US/08395507; Patent No. 5578456; GENERL INFORMATION: APPLICANT: Fulimura, Katsuya
                                                                                                                                                                                                                                                                                                                                                              NAME: BENT, Stephen A. REGISTRATION NUMBER: 29,768 REFERENCE/DOCKET NUMBER: 85
                                                                                                                                                                                                                                                                                                                                                                                                                                         TELECOMMUNICATION INFORMATION: TELEPHONE: (202)672-5300
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          TELEFAX: (202)672-5399
TELEX: 904136
INFORMATION FOR SEQ ID NO: 10:
SEQUENCE CHARACTERISTICS:
                                                                          Floppy disk
                                                                                                                                                                                                                                                                                                                                   ATTORNEY/AGENT INFORMATION:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         LENGTH: 472 amino acids
                                                                                                                                                                                                                                                                                                                                                                                                                                                       TELEPHONE: (202)01-
meterax: (202)672-5399
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 Conservative
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 ; MOLECULE TYPE: protein US-09-115-746-10
                                         COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    amino acid
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             linear
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     Best Local Similarity
                   20007-5109
                                                                                                                                                                                                                                  FILING DATE:
                                                                                                                                                                                                                                                                                                            FILING DATE:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           TOPOLOGY:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              Matches 255;
                                                                                                                                                       SOFTWARE:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 US-08-395-507-2
COUNTRY:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               Query Match
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  125
```

9

-- 325

```
283 ELALEKKSAPSPLTYRGTWMVREDGIVELSLVSSEQSKAPHEK------
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  TELEFAX: 215-875-8394
INFORMATION FOR SEQ ID NO: 4:
SEQUENCE CHARACTERISTICS:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   FILING DATE: 05-MAY-1994
ATTORNEY/AGENT INFORMATION:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        426 amino acids
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           Query Match 27.7%
Best Local Similarity 95.0%
Matches 228; Conservative
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      MOLECULE TYPE: protein
                                                                                                                                                                                                                                                                                                                                                      CORRESPONDENCE ADDRESS:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 COMPUTER READABLE FORM:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         REGISTRATION NUMBER:
                                                                                                                                                                                                                                                                                                                                                                                                        CITY: Philadelphia
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             APPLICATION NUMBER:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      APPLICATION NUMBER:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     OPERATING SYSTEM:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           amino acid
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     TOPOLOGY: linear
                                                                                                                                                                                                                           GENERAL INFORMATION:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             CLASSIFICATION:
                                                                                                                                                                                                                                                                                                                                                                                                                                                USA
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              FILING DATE:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                19102
                                                                                                                                                                                                                                                                                                                                                                                                                               ЬA
                                                                                                                                                                     US-08-737-248-4
                                                                                                                                                                                                                                                                                                                                                                                                                                                COUNTRY:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           US-08-737-248-4
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        LENGTH:
                                                                                                                                                                                                                                                                                                                                                                                                                             STATE:
                                                                                                                                                    RESULT 10
                                                         δò
                                                                                              q
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              qq
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      8
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           185 IEAIPQIDKYLKSSKYIAWPLQGWQATFGGGDHPPKSDHLVPRGSML----RLLKRNIQI 240
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          298 NIGANSFSAETPVD-----LKKENETELANIS-----GPHKKSTSTSTRKRARSSKKKA 346
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        65 VKLTQSMAIIRYIADKHNMLGGCPKERAEISMLEGAVLDIRYGVSRIAYSKDFETLKVDF 124
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             125 LSKLPEMLKMFEDRLCHKTYLNGDHVTHPDFMLYDALDVVLYMDPMCLDAFPKLVCFKKR 184
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          241 SKRIVFTILKQKAFKGNHP---CVPSVCTITYSRFHCLPDTLKSLLPMSSKTTLSMLPQV 297
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  -----ADGTAQKV 282
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  65; Indels 69; Gaps
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    3 PILGYWKIKGLVQPTRLLLEYLEEKYEEHLYERDEGDKWRNKKFELGLEFPULPYYIDGD 62
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   5 PILGYWKIKGLVQPTRLLLEYLEEKYEEHLYERDEGDKWRNKKFELGLEFPNLPYYIDGD 64
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              27.8%; Score 1207.5; DB 1; Length 362; 61.6%; Pred. No. 1.7e-90;
                                                   Anti-treponema Pallidum Antibody Immunoassay
                                                                                                                                                                                                                                                                                                        SOFTWARE: Patentin Release #1.0, Version #1.30 CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/395,507
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              242 KAEKVECALKGGIFRGTLPAADCPGIDTTVTFN-------
                                                                                                                           Pollock, Vande Sande & Priddy
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  24; Mismatches
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          , LUCATION: 1..362
; OTHER INFORMATION: /note= "G17 Antigen"
US-08-395-507-2
                                                                                                                                          1990 M Street, NW, Suite 800
                                                                                                                                                                                                                                                                                                                                                                                                PRIOR APPLICATION DATA:
APPLICATION NUMBER: JP Hei-6-54672
FILLING DATE: 28-FEB-1995
ATTORNEY/AGENT INPORMATION:
                                                                                                                                                                                                                                                                                                                                      FILING DATE: 27-FEB-1995
GLASSIFICATION: 436
                 Fujii, No. 5578456uyuki
Okada, Masahisa
                                                                                                                                                                                                                                                                       IBM PC compatible
YSTEM: PC-DOS/MS-DOS
                                                                                                                                                                                                                                                                                                                                                                                                                                                                         NAME: Amernick, Burton A.
REGISTRATION NUMBER: 24,852
REFERENCE/DOCKET NUMBER: 1581/
TELECOMMUNICATION INPORMATION:
TELEPHONE: (202) 331-7111
TELEPAS: (202) 293-6229
INFORMATION FOR SECION NO: 2:
SEQUENCE CHARACTERISTICS:
                                                                                                                                                                                                                                                        Floppy disk
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      LENGTH: 362 amino acids
TYPE: amino acid
STRANDEDNESS:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             not relevant
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               Best Local Similarity 61.6%
Matches 253; Conservative
Ueno, Eiichi
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   MOLECULE TYPE: protein
                                                                                                           CORRESPONDENCE ADDRESS:
                                                                                                                                                                                                                                    COMPUTER READABLE FORM:
                                                                                      NUMBER OF SEQUENCES:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    NAME/KEY: Protein
                                                  TITLE OF INVENTION:
                                                                                                                                                                                                                                                                         COMPUTER: IBM PC OPERATING SYSTEM:
                                                                                                                                                CITY: Washington STATE: Pr
                                                                                                                                                                                                      USA
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               ANTI-SENSE: NO FEATURE:
                                                                                                                                                                                                                                                        MEDIUM TYPE:
                                                                                                                                                                                                                     20036
                                                                                                                             ADDRESSEE:
APPLICANT:
                                                                                                                                                                                                      COUNTRY:
                                  APPLICANT:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                Query Match
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         q
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     qq
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                qq
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   ŏ
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          δ
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               δ
```

```
63 VKLTQSMAILKYIADKHNMLGGCPKERAEISMLEGAVLDIRYGVSRIAYSKDFETLKVDF 122
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               65 VKLTQSMALIRYIADKHNMLGGCPKERAEISMLEGAVEDIRYGVSRIAYSKDFETLKVDF 124
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   125 LSKLPEMLXMFEDRLCHKTYLNGDHVTHPDFMLYDALDVVLYMDPMCLDAFPKLVCFKKR 184
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                Indels 4; Gaps
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       5 PILGYWKIKGLVQPTRLLLEYLEEKYEEHI.YERDEGDKWRNKKFELGLEFPNLPYYIDGD 64
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 3 PILGYWKIKGLVQPTRLLLEYLEEKYEEHLYERDEGDKWRNKKFELGLEFPNLPYYIDGD 62
347 TDSVSDKIDESVASYDSSTHLRRSSRSKKPVNYNSSSESESEEQISKATKK 397
                                    326 --BLYBLIDSNSVRYMGAPGAGKPSKEMAPF-------YVLKKTKK 362
                                                                                                                                                                                                                                                                                        APPLICANT: Guemene, Daniel.
APPLICANT: Zadworny, David
APPLICANT: Karatzas, Costas
TITLE OF INVENTION: USE OF PROLACTIN FOR PREVENTING AND/OR
TITLE OF INVENTION: TREATING BIRD BROODINESS
NUMBER OF SEQUENCES: 23
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   27.7%; Score 1206; DB 3; Length 426; 95.0%; Pred. No. 2.8e-90; Live 2; Mismatches 6; Indels
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 SOFTWARE: Patentin Release #1.0, Version #1.30 CURRENT APPLICATION DATA:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               ADDRESSEE: WEISER & ASSOCIATES
STREET: 230 South Fifteenth Street, Suite 500
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           PRIOR APPLICATION DATA:
APPLICATION NUMBER: PCT/FR95/00576
FILING DATE: 03-MAY-1995
PRIOR APPLICATION DATA:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              REFERENCE/DOCKET NUMBER: 989.6411P
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           JMBER: US/08/737,248
28-APR-1997
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               PC-DOS/MS-DOS
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               FR 94/05550
                                                                                                                                                                                                ; Sequence 4, Application US/08737248
; Patent No. 6114305
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         /AGENT IN CO.
Weiser, Gerard J.
WHINER: 19,763
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       TELECOMMUNICATION INFORMATION: TELEPHONE: 215-875-8383
```

280 MLLSV 284

```
65 VKLTQSMAIIRYIADKHNMLGGCPKERAEISMLEGAVLDIRYGVSRIAYSKDFETLKVDF 124
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         G3 VKLTQSMAIIRYIADKHNMLGGCPKERAEISMLEGAVLDIRYGVSRIAYSKDFETLKVDF 122
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        125 LSKLPEMLKMFEDRLCHKTYLNGDHVTHPDFMLYDALDVVLYMDPMCLDAFPKLVCFKKR 184
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              185 IEAIPQIDKYLKSSKYIAWPLQGWQATFGGGDHPPKSDHLVPRGSMLRLLKRNIQISKRI 244
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 245 VFTILKQKAFKGNHPCVPSVCTITYSRFHCLP-----DTLKSLLPM-----SSKTTLS 292
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          35; Gaps
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        3 PILGYWKIKGLVQPTRLLLEYLEEKYEEHLYERDEGDKWRNKKFELGLEFPNLPYYIDGD 62
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  5 PILGYWKIKGLVQPTRLLLEYLEEKYEEHLYERDEGDKWRNKKFELGLEFPNLPYYIDGD 64
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            DB 2; Length 579;
                                                                                                                                                                                                                 APPLICANT: Elledge, Stephen J.
PSPLICANT: Liu, Olighua
TITLE OF INVENTION: Rapid Subcloning Using Site-Specific
TITLE OF INVENTION: Recombination
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       Indels
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    27.7%; Score 1205.5; DB 2;
78.0%; Pred. No. 5e-90;
... utematches 20;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               SOFTWARE: Patentin Release #1.0, Version #1.30 CURRENT APPLICATION DATA:
                                                                                                                                                                                                                                                                                                                                                                     220 Montgomery Street, Suite 2200
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           REGISTRATION NUMBER: 40,027
REGISTRATION NUMBER: 40,027
REPERENCE/POCKET NUMBER: BCM-02681
*TELECOMMUNICATION INFORMATION:
TELEPHONE: (415) 397-8338
INFORMATION FOR SEQ ID NO: 11:
SEQUENCE CHARACTERISTICS:
LENGTH: 579 antho acids
TYPE: amino acid
                                                                                                                                                                                                                                                                                                                                                                                     CITY: San Francisco
STATE: California
COUNTRY: United States of America
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            APPLICATION NUMBER: US/08/864,22
                                                                                                                                                                                                                                                                                                                                              Medlen & Carroll, LLP
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
                                                                                                                                             Sequence 11, Application US/08864224 Patent No. 5851808 GENERAL INFORMATION:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         ATTORNEY/AGENT INFORMATION:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          Query Match 27.7%;
Best Local Similarity 78.0%;
Matches 238; Conservative
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         MOLECULE TYPE: protein
                                                                                                                                                                                                                                                                                                      NUMBER OF SEQUENCES: 1
CORRESPONDENCE ADDRESS:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                            COMPUTER READABLE FORM:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         CLASSIFICATION: 435
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       inear
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      293 MLPQV 297
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     FILING DATE
                                                                                                                                                                                                                                                                                                                                                                                                                                    COUNTRY: U
                                                                                                                                                                                                                                                                                                                                                   ADDRESSEE:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   TOPOLOGY:
                                                                                                                                  US-08-864-224-11
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           US-08-864-224-11
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       ð
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              õ
```

```
TITLE OF INVENTION: Human Telomerase Catalytic Subunit
                                                                                                                                                                                                                                                                                                                                                                                                                                                           OFFWARE: Patentin Release #1.0, Version #1.30 CURRENT APPLICATION DATA:
                                                                                                                                                                                                                                                                       CORRESPONDENCE ADDRESS:
ADDRESSEE: Townsend and Townsend and Crew LLP
STREET: Two Embarcadero Center, Eighth Floor
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  015389-002610US
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       APPLICATION NUMBER: WO PCT/US97/17618
FILING DATE: 01-0CT-1997
PRIOR APPLICATION DATA:
APPLICATION NUMBER: WO PCT/US97/17885
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           IMBER: US/08/974,549A
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     ARTLING DATE: 14-AUG-1997
PRIOR APPLICATION DATA:
APPLICATION DATA:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    US 08/854,050
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               APPLICATION NUMBER: US 08/724,643 FILLING DATE: 01-0CT-1996
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          US 08/844,419
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   US 08/846,017
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          US 08/851,843
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             us 08/911,312
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    JMBER: US 08/912,951
14-AUG-1997
                                    ; Sequence 604, Application US/08974549A; Patent No. 6166178; GENERAL INFORMATION:
                                                                                                                                                                       APPLICANT: Morin, Gregg B.
APPLICANT: Harley, Calvin B.
APPLICANT: Andrews, William H.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              36,429
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 TELEPHONE: (415) 576-0200
TELEPAX: (415) 576-0300
INFORMATION FOR SEQ ID NO: 604:
                                                                                                                                                                                                                                                                                                                                                                                                                                                       IBM PC compatible
                                                                                         Cech, Thomas R.
Lingner, Joachim
Nakamura, Toru
Chapman, Karen B.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   TELECOMMUNICATION INFORMATION
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           NAME: Apple, Randolph Ted REGISTRATION NUMBER: 36,42
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         FILING DATE: 06-MAY-1997
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           FILING DATE: 14-AUG-1997
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       FILING DATE: 01-OCT-1997
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    25-APR-1997
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      09-MAY-1997
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         ATTORNEY/AGENT INFORMATION:
                                                                                                                                                                                                                                                                                                                                                                                                          COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        18-APR-1997
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  REFERENCE/DOCKET NUMBER:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    SEQUENCE CHARACTERISTICS:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          PRIOR APPLICATION DATA: APPLICATION NUMBER: 1
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          FILING DATE: 18-APR-1
PRIOR APPLICATION DATA:
APPLICATION NUMBER: U
FILING DATE: 25-APR-1
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           PRIOR APPLICATION DATA:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   PRIOR APPLICATION DATA: APPLICATION NUMBER:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  PRIOR APPLICATION DATA:
APPLICATION NUMBER:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         PRIOR APPLICATION DATA:
                                                                                                                                                                                                                                                                                                                                   CITY: San Francisco
STATE: California
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     APPLICATION NUMBER:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               APPLICATION NUMBER:
                                                                                                                                                                                                                                                     NUMBER OF SEQUENCES:
                                                                                                                                                                                                                                                                                                                                                                                          94111-3834
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     FILING DATE: 19
CLASSIFICATION:
                                                                                                                                                                                                                                                                                                                                                                          USA
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        FILING DATE:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             FILING DATE:
RESULT 12
US-08-974-549A-604
                                                                                                                                                                                                                                                                                                                                                                                                                                                       COMPUTER:
                                                                                                                                    APPLICANT:
                                                                                                                                                       APPLICANT:
                                                                                                APPLICANT:
                                                                                                                    APPLICANT:
                                                                                                                                                                                                                                                                                                                                                                          COUNTRY:
```

515 amino acids

LENGTH:

```
qq
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      123 LSKLPEMLKMFEDRLCHKTYLNGDHVTHPDFMLYDALDVVLYMDPMCLDAFPKLVCFKKR 182
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          185 IEAIPQIDKYLKSSKYIAWPLQGWQATFGGGDHPPKSDHLVPRGS-------M 230
                                                                                                                                                                                                                                                                                                                                                                                                                                                                 65 VKLTQSMAIIRYIADKHNMLGGCPKERAEISMLEGAVLDIRYGVSRIAYSKDFETLKVDF 124
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             125 LŞKLPEMLKMFEDRLCHKTYLNGDHVTHPDFMLYDALDVVLYMDPMCLDAFPKLVCFKKR 184
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         ---KRIVFTILKOKAFKGNHPCV 261
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          242 VELLRSFFYVTETTFQKNRLFFYRPSVWSKLQSIGIRQHLKRVQLRELSEAEVRQHREAR 301
                                                                                                                                                                                                                                                                                                                                                       55; Gaps
                                                                                                                                                                                                                                                                                                                                                                                                           5 PILGYWKIKGLVQPTRLLLEYLEEKYEEHLYERDEGDKWRNKKFELGLEFPNLPYYIDGD 64
                                                                                                                                                                                                                                                                                                            27.7%; Score 1204.5; DB 4; Length 515; 69.7%; Pred. No. 5e-90; Live 17; Mismatches 33; Indels 55;
                                                                                                                                                                         thrombin cleavage sequence, recognition sequence for heart muscle protein
                                                                                                                                                                                                               kinase, residues introduced by cloning and hTRT protein fragment"
                                                                                                                                                   glutathione-S-transferase sequence
                                                                                                                                      /note= "fusion protein composed of
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                262 PSVCTITYSRFHCLPDTLKSLLPMSSKTTLSMLPQVNIGANSFSAE 307
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      Human Telomerase Catalytic Subunit
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         'SOFTWARE: Patentin Release #1.0, Version #1.30 CURRENT APPLICATION DATA:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               Townsend and Townsend and Crew LLP
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 Two Embarcadero Center, Eighth Floor
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            APPLICATION NUMBER: US/08/974,549A
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                Sequence 605, Application US/08974549A Patent No. 6166178 GENERAL INFORMATION:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         PC-DOS/MS-DOS
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            APPLICANT: Cech, Thomas R.
APPLICANT: Lingner, Joachim
APPLICANT: Nakamura, Toru
APPLICANT: Chapman, Karen B.
APPLICANT: Morin, Gregg B.
APPLICANT: Harley, Calvin B.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  Andrews, William H.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     231 LRLLKRNIQIS-----
                                                                                                                                                                                                                                                                                                                                                   Conservative
                                                 MOLECULE TYPE: protein
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          CORRESPONDENCE ADDRESS:
                                                                                                                              OTHER INFORMATION:
OTHER INFORMATION:
OTHER INFORMATION:
OTHER INFORMATION:
OTHER INFORMATION:
OTHER INFORMATION:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  San Francisco
                                                                                               NAME/KEY: Protein
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 TITLE OF INVENTION:
NUMBER OF SEQUENCES:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         California
amino acid
                                     TOPOLOGY: linear
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         OPERATING SYSTEM:
                                                                                                                                                                                                                                                                                                                               Local Similarity
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            94111-3834
            STRANDEDNESS:
                                                                                                                                                                                                                                                    US-08-974-549A-604
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 US-08-974-549A-605
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             ADDRESSEE:
                                                                                                             LOCATION:
                                                                                                                                                                                                                                                                                                                                                 Matches 241;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  APPLICANT:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         COUNTRY:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 STREET:
                                                                                                                                                                                                                                                                                                                Query Match
                                                                                                                                                                                                                                                                                                                                                                                                                              g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                   οy
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        q
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                ô
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   qq
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                ò
```

```
65 VKLTQSMALIRYIADKHNMLGGCPKERAEISMLEGAVLDIRYGVSRIAYSKDFETLKVDF 124
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             48; Gaps
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              5 PILGYWKIKGLVQPTRLLLEYLEEKYEEHLYERDEGDKWRNKKFELGLEFPNLPYYIDGD 64
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    3 PILGYWKIKGLVQPTRLLLEYLEEKYEEHLYERDEGDKWRNKKFELGLEFPNLPYYIDGD 62
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          thrombin cleavage sequence, recognition sequence for heart muscle protein kinase, residues introduced by cloning and hTRT protein fragment.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    27.7%; Score 1203; DB 4; Length 514; 71.6%; Pred. No. 6.6e-90;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             Indels
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        glutathione-S-transferase sequence,
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       /note= "fusion protein composed of
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           36;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                015389-002610US
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           13; Mismatches
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                WO PCT/US97/17618
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      WO PCT/US97/17885
                                                                                                                                                                                                                                                                                                         ...ING DATE: 14-AUG 1997
PRIOR APPLICATION DATA:
FILTMC
                                                                                                          JMBER: US 08/844,419
18-APR-1997
                                                        US 08/724,643
                                                                                                                                                                   US 08/846,017
                                                                                                                                                                                                                     APPLICATION NUMBER: US 08/851,843
                                                                                                                                                                                                                                                                            US 08/854,050
                                                                                                                                                                                                                                                                                                                                                                                                                                           us 08/915,503
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              36,429
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            REFERENCE/DOCKET NUMBER: 01
TELECOMMUNICATION INFORMATION:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             (415) 576-0200
(415) 576-0300
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         NAME: Apple, Randolph Ted REGISTRATION NUMBER: 36,4
19-NOV-1997
                                                                      01.-OCT-1996
                                                                                                                                                                                FILING DATE: 25-APR-1997
PRIOR APPLICATION DATA:
                                                                                                                                                                                                                                     FILING DATE: 06-MAY-1997 PRIOR APPLICATION DATA:
                                                                                                                                                                                                                                                                                          FILING DATE: 09-MAY-1997
PRIOR APPLICATION DATA:
                                                                                                                                                                                                                                                                                                                                                                                                    FILING DATE: 14-AUG-1997
PRIOR APPLICATION DATA:
                                                                                                                                                                                                                                                                                                                                                                                                                                                         14-AUG-1997
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    FILING DATE: 01-OCT-1997
ATTORNEY/AGENT INFORMATION:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 01-0CT-1997
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       : 514 amino acids
amino acid
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                INFORMATION FOR SEQ ID NO:
SEQUENCE CHARACTERISTICS:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       Matches 244; Conservative
                                 PRIOR APPLICATION DATA;
                                                                                       PRIOR APPLICATION DATA: APPLICATION NUMBER: 1
                                                                                                                                          PRIOR APPLICATION DATA:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                            PRIOR APPLICATION DATA:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   PRIOR APPLICATION DATA:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           MOLECULE TYPE: protein
                                                                                                                                                                                                                                                                                                                                                                                                                                           APPLICATION NUMBER:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              APPLICATION NUMBER:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  APPLICATION NUMBER:
                                                    APPLICATION NUMBER:
                                                                                                                                                                 APPLICATION NUMBER:
                                                                                                                                                                                                                                                                            APPLICATION NUMBER:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  NAME/KEY: Protein
LOCATION: 1..515
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       OTHER INFORMATION:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           OTHER INFORMATION:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             OTHER INFORMATION: OTHER INFORMATION:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        OTHER INFORMATION:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            INFORMATION:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               inear.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        Best Local Similarity
                 CLASSIFICATION:
                                                                      FILING DATE:
                                                                                                                            FILING DATE:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               FILING DATE:
                                                                                                                                                                                                                                                                                                                                                                                                                                                         FILING DATE:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      STRANDEDNESS:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  US-08-974-549A-605
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   TELEPHONE:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               TOPOLOGY:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  TELEFAX:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       LENGTH:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      Query Match
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            OTHER
```

7;

```
APPLICATION NUMBER: US 08/915,503 FILING DATE: 14-AUG-1997 PRIOR APPLICATION DATA:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                PRIOR APPLICATION DATA:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        PRIOR APPLICATION DATA:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               FILING DATE: 14-AUG-1
PRIOR APPLICATION DATA:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            PRIOR APPLICATION DATA: APPLICATION NUMBER:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         PRIOR APPLICATION DATA: APPLICATION NUMBER:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       COMPUTER READABLE FORM:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           APPLICATION NUMBER:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      APPLICATION NUMBER:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      94111-3834
                                                                                                                                                                                                                                                                                                                                                                                                 GENERAL INFORMATION:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      USA
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                FILING DATE:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            FILING DATE:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            FILING DATE:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         FILING DATE:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           MEDIUM TYPE:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          FILING DATE:
                                                                                                                                                                                                                                                                                                                                            US-08-974-549A-603
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               ADDRESSEE:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 COMPUTER:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   APPLICANT:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    COUNTRY:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              STREET:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    STATE:
                                                                                                                                                                                                                                                                                                                        RESULT 15
                                                                                                                                                                                                                                                               qq
                                                                                                                                                  ò
                                                                                                                                                                                       a
                                                                                                                                                                                                                             ŏ
                                   P
                                                                           δ
                                                                                                            q
                   183 IEAIPQIDKYLKSSKYIAWPLQGWQATFGGGDHPPKSD-LVPRGS------RRAS 230
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              Gaps
185 IEAIPQIDKYLKSSKYIAWPLQGWQATFGGGDHPPKSDHLVPRGSMLRLLKRNIQISKRI 244
                                                                                                              231 VGSVTKATSLEGALSGTRHSHPSVGRQHHAGPPSTSRPPRPWDTPCPPVYAETKHFLYSS 290
                                                                           -- PSV------CTITYSRF-HCL---
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            ij
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   27.6%; Score 1201.5; DB 1; Length 352; 99.1%; Pred. No. 5e-90; Live 0; Mismatches 1; Indels 1;
                                                                                                                                                                                                                                                                                                                                                                                                                       TITLE OF INVENTION: Anti-treponema Pallidum Antibody TITLE OF INVENTION: Immunoassay
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    SOFTWARE STATEM 1. 10. Version #1.30 CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/395,507
FILING DATE: 27-FEB-1995
CLASSIFICATION: 436
PRIOR APPLICATION DATA:
APPLICATION NUMBER: JP Hei-6-54672
FILING DATE: 28-FEB-1995
ATTORNEY/AGENT INFORMATION:
                                                                                                                                                      276 -----PDTLKSLLPMSSKTTLSMLPQVNIGANSFSAETP 309
                                                                                                                                                                                       291 GDKEQLRPSFLLSSLRPSLTGARRLVETIFLGSRPWMPGTP 331
                                                                                                                                                                                                                                                                                                                                                                                                                                                          NUMBER OF SEQUENCES: 2
CORRESPONDENCE ADDRESS:
ADDRESSEE: Pollock, Vande Sande & Priddy
STREET: 1990 M Street, NW, Suite 800
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             /note= "G15 Antigen"
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              NAME: Amernick, Burton A. REGISTRATION NUMBER: 24,852
REFERENCE/DOCKET NUMBER: 1581/007
TELECOMMUNICATION INFORMATION:
                                                                                                                                                                                                                                                                                                                                            APPLICANT: Fujimura, Katsuya
APPLICANT: Ueno, Biichi
APPLICANT: Fujii, No. 5578456uyuki
APPLICANT: Okada, Masahisa
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                COMPUTER: IBM PC compatible OPERATING SYSTEM: PC-DOS/MS-DOS
                                                                             245 VFTILKOKAFKG-----NHPCV---
                                                                                                                                                                                                                                                                                    Sequence 1, Application US/08395507
Patent No. 5578456
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   TELEPHONE: (202) 331-7111
TELEFAX: (202) 293-6229
INFORMATION FOR SEQ ID NO: 1:
SEQUENCE CHARACTERISTICS:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    Floppy disk
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  : 352 amino acids
amino acid
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       Query Match
Best Local Similarity 99.15
Matches 225; Conservative
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           not relevant
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       MOLECULE TYPE: protein HYPOTHETICAL: NO
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 COMPUTER READABLE FORM:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            ; LOCATION: 1..352
; OTHER INFORMATION:
US-08-395-507-1
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            NAME/KEY: Protein
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        ORGANISM: rabbit
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         Washington
                                                                                                                                                                                                                                                                                                                              GENERAL INFORMATION:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   2
Z
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               USA
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   ORIGINAL SOURCE
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        STRANDEDNESS:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  MEDIUM TYPE:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 HYPOTHETICAL:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 20036
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               ပ္က
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    LENGTH:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               COUNTRY:
                                                                                                                                                                                                                                                       RESULT 14
US-08-395-507-1
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           STATE:
                                                                                                                                                                                         q
                                                                                                                qq
                                                                                                                                                        ò
                                         g
                                                                               οy
```

```
123 LSKLPEMLKMFEDRLCHKTYLNGDHVTHPDFMLYDALDVVLYMDPMCLDAFPKLVCFKKR 182
                                                                                    65 VKLTQSMAIIRYIADKHNMLGGCPKERAEISMLEGAVLDIRYGVSRIAYSKDFETLKVDF 124
                                                                                                                                                                        125 LSKLPEMLKMFEDRLCHKTYLNGDHVTHPDFMLYDALDVVLYMDPMCLDAFPKLVCFKKR 184
                    5 PILGYWKIKGLVQPTRLILEYLEEKYEEHLYERDEGDKWRNKKFELGLEFPNLPYYIDGD 64
                                                                                                                                                                                                                                                                                      183 IEAIPQIDKYLKSSKYIAWPLQCWQATFGGGDHPPKSD-LVPRGSEL 228
                                                                                                                                                                                                                                                             185 IEAIPQIDKYLKSSKYIAWPLQGWQATFGGGDHPPKSDHLVPRGSML 231
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            TITLE OF INVENTION: Human Telomerase Catalytic Subunit NUMBER OF SEQUENCES: 727 CORRESPONDENCE ADDRESS:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     SOFTWARE: PatentIn Release #1.0, Version #1.30 CURRENT APPLICATION DATA: APPLICATION NUMBER: US/08/974.5444
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     SEE: Townsend and Townsend and Crew LLP: Two Embarcadero Center, Eighth Floor San Francisco California
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 JMBER: US/08/974,549A
19-NOV-1997
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                CLASSIFICATION: 536
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/724,643
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           US 08/911,312
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            US 08/844,419
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            US 08/851,843
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            WUMBER: US 08/854,050
09-MAY-1997
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     JMBER: US 08/912,951
14-AUG-1997
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         US 08/846,017
                                                                                                                                                                                                                                                                                                                                                                                                                 Sequence 603, Application US/08974549A Patent No. 6166178
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   APPLICANT: Cech, Thomas R.
APPLICANT: Lingner, Joachim
APPLICANT: Nakamura, Toru
APPLICANT: Chapman, Karen B.
APPLICANT: Morin, Gregg B.
APPLICANT: Harley, Calvin B.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        Andrews, William H.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              E: Floppy disk
IBM PC compatible
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            FILING DATE: 01-007-1996
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              FILING DATE: 25-APR-1997
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              14-AUG-1997
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     18-APR-1997
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  06-MAY-1997
```

```
230 -----KRIVFTILK 250
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         65 VKLTQSMAIIRYIADKHNMLGGCPKERAEISMLEGAVLDIRYGVSRIAYSKDFETLKVDF 124
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             LSKLPEMLKMFEDRICHKTYLNGDHVTHPDFMLYDALDVVLYMDPMCLDAFPKLVCFKKR 184
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         242 HGSVTKMSVYVVELLRSFFYVTETTFQKNRLFFYRPSVWSKLQSIGIRQHLKRVQLRELS 301
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              311 D-------LKKENE-----TELANISGPHKKSTSTSTRKRARSSKKK----- 345
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     351 ERLTSRVKALFSVLNYERARRPGLLGASVLGLDDIHRAWRTFVLRVRAQDPPPELYFVKV 410
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      27.6%; Score 1200.5; DB 4; Length 530;
50.4%; Pred. No. 1.1e-89;
Live 49; Mismatches 91; Indels 119; Gaps
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       251 QKAFKGNHPCVPSVCTITYSRFHCLPDTLKSLLPMSSKTTLSMLPQVNIGANSFSAETPV 310
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               346 ----ATDSV-SDKIDESVASY------DSSTHLRRSSRSKKPVNYNSSS 383
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                5 PILGYWKIKGLVOPTRLLLEYLEEKYEEHLYERDEGDKWRNKKFELGLEFPNLPYYIDGD 64
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      thrombin cleavage sequence, recognition sequence for heart muscle protein Kinase, residues introduced by cloning, eight consecutive His residues and hTRT protein fragment."
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  185 IEAIPQIDKYLKSSKYIAWPLQGWQATFGGGDHPPKSDHLVPRGS------
                                                                                                                                                                                                                                                                                                                                                                                                                            /note= "fusion protein composed of
glutathione-S-transferase sequence
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      384 ESESEEQISKATKKVKQKEEEEYVEEVDEKSLKNESSSDEFE 425
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  465 LTDLQPYMRQFVAHLQETSPLRDAVVIEQSSSLNEASSGLFD 506
                                                                                                            NAME: Apple, Randolph Ted
REGISTRATION NUMBER: 36,429
REFERENCE/DOCKET NUMBER: 015389-002610US
TELECOMMUNICATION INFORMATION:
TELEPHONE: (415) 576-0200
INFORMATION FOR SEQ ID NO: 603:
SEQUENCE CHARACTERISTICS:
APPLICATION NUMBER: WO PCT/US97/17618
FILING DATE: 01-0CT-1997
PRIOR APPLICATION DATA:
APPLICATION NUMBER: WO PCT/US97/17885
FILING DATE: 01-0CT-1997
                                                                                             ATTORNEY/AGENT INFORMATION:
                                                                                                                                                                                                                                                                       : 530 amino acids
amino acid
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           Matches 263; Conservative
                                                                                                                                                                                                                                                                                                                                              MOLECULE TYPE: protein
                                                                                                                                                                                                                                                                                                                                                                                                        LOCATION: 1..530
OTHER INFORMATION:
                                                                                                                                                                                                                                                                                                                                                                                     NAME/KEY: Protein
                                                                                                                                                                                                                                                                                                                            linear
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        Best Local Similarity
                                                                                                                                                                                                                                                                                                          STRANDEDNESS:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               US-08-974-549A-603
                                                                                                                                                                                                                                                                                                                                TOPOLOGY:
                                                                                                                                                                                                                                                                         LENGTH:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        Query Match
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              qq
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           οy
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       qq
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              δ
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 qq
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       Ω
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            QQ
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      qq
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                ò
```

Search completed: January 15, 2002, 13:50:57 Job time: 326 sec

```
GenCore version 4.5
Copyright (c) 1993 - 2000 Compugen Ltd.
```

OM protein - protein search, using sw model

January 15, 2002, 14:07:05; Search time 45.07 Seconds (without alignments) 301.812 Million cell updates/sec Run on:

1 DDHAPREMFDCLDKPIPWRG.......RRLTARKRRSRKEEVEEDEK 371 US-09-724-296-4 1979 Title: Perfect score: Sequence:

BLOSUM62 Gapop 10.0 , Gapext 0.5 Scoring table:

100059 segs, 36664827 residues Searched: 100059 Total number of hits satisfying chosen parameters:

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

SwissProt_39:* Database :

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Description			2					P31409 drosophila			mand				Q19626 caenorhabdi	baci	mus m	P21281 homo sapien		066875 aquifex aeo							_	P72241 providencia			ace	9	P55161 rattus norv
ID	YWJD_BACSU	VATB_YEAST	VATB_DESSY	VATB_ARATH	VATB_CYACA	VATB_ARCFU	VATB_THESI	VATB_DROME	GOG4_HUMAN	VATB_HELVI	VATB_MANSE	VAB1_BOVIN	VATB_CHICK	VATB_PYRAB	VATB_CAEEL	CYAA_BACAN	PROD_MOUSE	VAB2_HUMAN	VAB2_MOUSE	BIOF_AQUAE	VAB1_HUMAN	Y133_METJA	VAB1_GOSHI	VAB2_GOSHI	VAB2_BOVIN	VATB_PYRHO	VATB_CANTR	GCPE_PROST	RRPL_DUGBV	VAB2_ACEAT	VAB1_ACEAT	õ	NCP1_RAT
ength DB	320 1				500 1						494 1	513 1	453 1	465 1	491 1	800 1	497 1	511 1	511 1	373 1	513 1	273 1	488 1	386 1	511 1	465 1	511 1	365 1	4036 1	492 1	492 1	648 1	1128 1
% Query Match Length					5.5										5.1		5.1	5.1		5.1		Ŋ	ď.	4	4.9	4.9	4.9	4.8	4.8	4.8	4.7	4.7	4.7
Score	360	111.5	110	108.5	108	104.5	104	104	103.5	103	103	102	101.5	101.5	101	101	100.5	100.5	100.5	100	100	99.2	98.2	97.5	97.5		96	95	95	4		æ.	ë.
Result No.	1	7	m	4	Z.	9	7	80	6	10	11	12	13	14	15	16	17	18	19	20	21	22	23	24	25	56	27	28	29	30	31	32	33

O27035 methanobact	Q9ulw0 homo sapien	P35609 homo sapien	Q40079 hordeum vul	P50533 xenopus lae	P51835 bacillus su	Q05654 schizosacch	Q9y2a7 homo sapien	059258 pyrococcus	Q90734 gallus gall	P42694 homo sapien	Q9hm64 thermoplasm
VATB_METTH	DIL2_HUMAN	AAC2_HUMAN	VAB2_HORVU	XCPE_XENLA	FTSY_BACSU	RDPO_SCHPO	NCP1_HUMAN	UPPS_PYRHO	AAC4_CHICK	Y054_HUMAN	VATB_THEAC
Н		7	_	Н	1	-	<u>,</u>	-	٦	П	1
463	747	894	483	1203	329	1333	1128	264	904	1942	460
4.7	4.7	4.7	4.7	4.7	4.6	4.6	4.6	4.6	4.6	4.6	4.6
	93	63	2.5	12.5	92	92	91.5	91	91	91	90.5
93			0								

ALIGNMENTS

```
Db 257 RSHADYVDANFLLPLLERFRQWGTNIDFMIEAKQKDKALLRL 298
```

```
This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation - the Buropean Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See http://www.isb-sib.ch/announce/or send an email to license@isb-sib.ch).
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       SUBUNIT: V-ATPASE IS AN HETEROMULTIMERIC ENZYME COMPOSED OF A PERIPHERAL CATALYTIC VI COMPLEX (MAIN COMPONENTS: SUBUNITS A, B, C, D, E, AND F) ATTACHED TO AN INTEGRAL MEMBRANE VO PROTON PORE COMPLEX (MAIN COMPONENT: THE PROTECLIPID PROTEIN).

SIMILARITY: BELONGS TO THE ATPASE ALPHA/BETA CHAINS FAMILY.
CAUTION: REF. 4 SEQUENCE WAS INCORRECTLY THOUGHT TO ORIGINATE FROM
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            "The sequence of 29.7 kb from the right arm of chromosome II reveals
                                                                                                                                                                                                                                                                                  Nelson H., Mandiyan S., Nelson N.;
"A conserved gene encoding the 57-kDa subunit of the yeast vacuolar
                                          P16140; P32123;
01-APR-1990 (Rel. 14, Created)
01-APR-1994 (Rel. 30, Last sequence update)
20-AUG-2001 (Rel. 40, Last annotation update)
VACUOLAR ATP SYNTHASE SUBUNIT B (EC 3.6.1.34) (V-ATPASE B SUBUNIT)
VACUOLAR PROTON PUMP B SUBUNIT) (V-ATPASE 57 KDA SUBUNIT)
VMAZ OR VATZ OR YBR127C OR YBR1002.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      Becam A.-M., Cullin C., Grzybowska E., Lacroute F., Nasr F.,
Ozier-Kalogeropoulos O., Palucha A., Slonimski P.P., Zagulski M.,
                                                                                                                                                              Saccharomyces cerevisiae (Baker's yeast).
Eukaryota: Fungi: Ascomycota: Saccharomycotina; Saccharomycetes;
Saccharomycetales; Saccharomycetaceae; Saccharomyces.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          complete open reading frames, of which ten correspond to new
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        MEDLINE=92156101; PubMed=1371275; Puopolo K., Kumamoto C., Adachi I., Mager R., Forgac M.; "Differential expression of the 'B' subunit of the vacuolar
                                   PRT;
                                                                                                                                                                                                                                                                                                                                                                                      Nelson H., Mandiyan S., Nelson N.;
J. Biol. Chem. 264:5313-5313(1989).
                                                                                                                                                                                                                                                                                                                                    Biol. Chem. 264:1775-1778(1989).
                                                                                                                                                                                                                                                                      MEDLINE=89109197; PubMed=2521486;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                        MEDLINE=94378717; PubMed=8091856;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        SEQUENCE OF 103-517 FROM N.A.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                EMBL; J04450; AAA66890.1; -.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  EMBL; X75891; CAA53486.1; -. EMBL; Z35996; CAA85084.1; -. EMBL; M83130; AAA30389.1; -.
                                 STANDARD;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            // (1994)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    PIR; S45996; S45996.
                                                                                                                                                                                                                                                  SEQUENCE FROM N.A.
                                                                                                                                                                                                                                                                                                                                                                                                                                        SEQUENCE FROM N.A.
                                                                                                                                                                                                                  NCBI_TaxID=4932;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          Herbert C.J.;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                VESICLES.
                                                                                                                                                                                                                                                                                                                                                                                                                                                        STRAIN=S288C
                                                                                                                                                                                                                                                                                                                      H+-ATPase."
                              VATB_YEAST
                                                                                                                                                                                                                                                                                                                                                                        REVISIONS.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  BOVINE.
              VATB_YEAST
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            genes
RESULT
```

```
18;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  207 NFSIVFAAMGVNL-----ETARFFKODFEENGSLERTSLFLNLANDPTIERIITP 256
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             257 RLALTTAEYLAYQTERHVLTILTDMSSYADALREVSAAREEVPGRRGYPGYMYTDLSTIY 316
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              419 VVGEEALSIEDK-----LSLEFLEKFEKTFITQGAYEDRTVFESLDQAWSLLR 466
                                                                                                                                                                                                                                                                                                                                                                                                                 161 MNSIARGOKIPIFSASGLPHNEIAAQICROAGLVRP-----TKDVHD-----GHEE 206
                                                                                                                                                                                                                                                                                                                                                                                                                                                     77 NFGIHFMRVSSDLFPFASHAKYGYTLEFAQSHLEEVG-----KLANK-----YNH 121
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             122 RLTMHPGQYTQIASPREVVV-----DSAIRDL-AYHDEILSRMKLNEQLNKDAVLII 172
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  173 HLGGTFECKKETLDRFRKNYQRLSDSVKARLVLENDDVSWSVQDLLPLCQELNIPLVLDW 232
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     317 ERAGRVEGRNGSITQI------PILTMPNDDITHPIPDLTGYITEGQIFVDRQL 364
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         233 HHHNIVPGTLREGSLDLMPLIPTIRETWTRKGITQKQH-----YS---ESADPTA 279
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  365 HNKGIYP-----PINVLPSLSRLMKSAIGEGMTRKDHGDVSNQLYAKYAIGKDAAAMKA 418
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               280 ISGMKRRAHSDRVFDFPPCDPTMDLMIEAKEK------EQAVFE-----LCR 320
                                                                                                                                                                                                                                                                                                                                             62; Mismatches 140; Indels 125; Gaps
                                                                                                                                                                                                                                                                                                                                                                              27 LNTILRSMKERVF-----CSRTCRITTIQRDGLESVKQLGTQNVLDLIKLVEWNH 76
                                                                                                                                                                             V -> L (IN REF. 1).
Q -> R (IN REF. 4).
DTRSSGKKKDASQEESLI -> TQEAPVRRTPAKKNL
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  20-AUG-2001 (Rel. 40, Created)
20-AUG-2001 (Rel. 40, Last sequence update)
20-AUG-2001 (Rel. 40, Last annotation update)
V-TYPE ATP SYNTHASE BETA CHAIN (EC 3.6.1.34) (V-TYPE ATPASE SUBUNIT
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    Shibui H., Hamamoto T., Yohda M., Kagawa Y.;
"The stabilizing residues and the functional domains in the hyperthermophilic V-Mrpase of Desulfurococcus.";
Hiochem. Biophys. Res. Commun. 234:341-345(1997).
-i. FUNCTION: PRODUCES ATP FROM ADP IN THE PRESENCE OF A PROTON GRADIENT ACKOSS THE MEMBRANE. THE ARCHAEAL BETA CHAIN IS A
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 Archaea; Crenarchaeota; Desulfurococcales; Desulfurococcaceae; Desulfurococcus.
                                                                                                                                                                                                                                                                                                            DB 1; Length 517;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      321 RY--ELQNPPCPLEIMGPEYDQTRD--GYYPPGAEKRLTARKRRSRKEE 365
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         467 IYPKEMLNRISP-KILDEFYDRARDDADEDEEDPDTRSSGKKKDASQEE 514
                                                                                                                                                                                                                                                    02A7DEC571EFF7C2 CRC64;
                                                                                                                                                            Hydrolase; ATP synthesis; Hydrogen ion transport.
                                                                                                                                                                                                                                                                                                        5.6%; Score 111.5; Di 20.0%; Pred. No. 0.26;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     463 AA
                                                                                                                 Pfam; PF01038; ATP_synt_B_c; 1.
PROSITE; PS00152; ATPASE_ALPHA_BETA; 1.
                                                                                  IPR000194; ATPasê_alpha_beta.
                                                                                                                                                                                                                                     (IN REF.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     PRT;
                                                                     InterPro; IPR003255; ATP_synt_B_c.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          MEDLINE=97320421; PubMed=9177272;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              Desulfurococcus sp. (strain SY).
                                                                                                        Pfam; PF00006; ATP-synt_ab; 1.
                                                                                                                                                                                                                                                    57749 MW;
                                                                                                                                                                                                                                                                                                                                               Conservative
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     STANDARD;
                                                                                                                                                                                                227
                                                                                                                                                                                 4
B42254; B42254.
               S46570; S46570.
                                                  S0000331; VMA2.
                                                                                                                                                                                                                                                  517 AA;
                                                                                                                                                                                                                                                                                                                            Local Similarity
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       SEQUENCE FROM N.A.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      NCBI_Tax1D=59822;
                                                                                                                                                                                              227
500
                                                                                                                                                                             79
                                8560;
                                                                                                                                                                                                                                                                                                                                         82;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   VATB_DESSY
                                                                                    InterPro;
                                                                                                                                                                                                                                                    SEQUENCE
                                                                                                                                                                               CONFLICT
                                                                                                                                                                                                CONFLICT
                                                                                                                                                                                                                CONFLICT
                                                                                                                                                                                                                                                                                                          Query Match
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   006505;
                              YEPD;
                                                    SGD;
                                                                                                                                                                                                                                                                                                                                           Matches
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  VATB_DESSY
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               ATPB.
                                                                                                                                                                                                                                                                                                                            Best
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 RESULT
DR
DR
DR
DR
FT
FT
                                                                                                                                                                                                                                                                                                                                                                                ŏλ
                                                                                                                                                                                                                                                                                                                                                                                                                 q
                                                                                                                                                                                                                                                                                                                                                                                                                                                       ô
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        q
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             Ω
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               q
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     q
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         õ
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               QΥ
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     2
```

```
δŏ
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    qq
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    δý
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    q
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    ŏ
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                Οp
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   ó
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                q
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                QQ
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 \overset{\circ}{\circ}
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                AC AC DIT
                                                                                                                                                                                                                                                                                                                                                                          13;
                                               This SWISS-PROT entry is copyright. It is produced through a collaboration
                                                                                                                           entitles requires a license agreement (See http://www.isb-sib.ch/announce/
or send an email to license@isb-sib.ch).
                                                                                                              Usage by and for commercial
                                                             between the Swiss Institute of Bioinformatics and the EMBL outstation the European Bioinformatics Institute. There are no restrictions on it use by non-profit institutions as long as its content is in no wa modified and this statement is not removed. Usage by and for commercia
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        129 QYTQIASPREVVV-----DSAIRDL-AYHDEILSRMKLNEQLNKDAVL;IIHLGGTFE 179
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       180 GKKETLDRFRKNYQRLSDSVKARLVLENDDVSWSVQDLLPLCQELNIPLVLDWHHHNIVP 239
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  342 -----PIDVLPSLSRLMKDGIGKGRTREDHPQLAQQLYAAYAEGRSLRDLVAVVGEEAL 395
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    396 SETDRKYLKFADRFEREFIAQRYDEDRSIFETLDLGWELL-----AELPESELKRVRKEY 450
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; Rosidae; eurosids II; Brassicales; Brassicaceae; Arabidopsis.
                                                                                                                                                                                                                                                                                                                                                                                                                                         ----- EQFAVVF---- 182
                                                                                                                                                                                                                                                                                                                                                                                                                                                                         87 SDLFPFASHAKYGYTLE---FAQSHLEEVGKLANK-------YNHRLTMHPG 128
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        183 -----AAMGITYEEANFFKKSFEETGAIERAVLFLNLADDPAIERIITPRMALTVA 233
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      234 EYLAFDYDMOVLVILTDMTNYAEALREISAAREEVPGRRGYPGYMYTDLATIYERGGRVR 293
                                                                                                                                                                                                                                                                                                                                                                          92; Gaps
                                                                                                                                                                                                                                                                                                                                                                                                          27 LNTILRSMKERVFCSRTCRITTIQRDGLESVKQLGTQNVLDLIKLVEWNHNFGIHFMRVS 86
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               VATE_ARATH STANDARD; PRT; 492 AA.
191574; 042205;
01-0CT-1989 (Rel. 12, Last sequence update)
01-0CT-1989 (Rel. 140, Last annotation update)
VACUOLAR ATP SYNTHASE SUBUNIT B (EC 3.6.1.34) (V-ATPASE B SUBUNIT)
(VACUOLAR APP SYNTHASE SUBUNIT) (V-ATPASE 57 KDA SUBUNIT).
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         240 GTLREGSLDLMPLIPTIRETWTRKGITQKQH------YSES---ADPTAISGMKRR
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    287 AHSDRVFDFPPCDPTMDLMIEAKEKEQAVFE-LCRRYELQNPPCPLEIMGPEYDQTRDGY
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                SEQUENCE FROM N.A. MEDLINE-89053962; PubMed-2903860; Manolson M.F., Ougllette B.F.F., Filion M., Poole R.J.; Manolson M.F., and homologies of the '57-kDa' nucleotide-binding
               SIMILARITY: BELONGS TO THE ATPASE ALPHA/BETA CHAINS FAMILY
                                                                                                                                                                                                                                                                                                                                        5.6%; Score 110; DB 1; Length 463; 19.1%; Pred. No. 0.29;
                                                                                                                                                                                                                                                                                                                                                                         55; Mismatches 153; Indels
                                                                                                                                                                                                                                                                             Hydrolase; ATP synthesis; Hydrogen ion transport.
SEQUENCE 463 AA; 52051 MW; 8D94195D02D50D9A CRC64;
                                                                                                                                                                                                                                                                                                                                                                                                                                         139 MNTLVRGOKLPIFSGSGLPHNMLAAQIARQAKVLGEE-----
                                                                                                                                                                                                                     Pfam: PF00006; ATP-synt_ab; 1.
Pfam: PF01038; ATP-synt_B_c; 1.
PROSITE; PS00152; ATPASE_ALPHA_BETA; 1.
                                                                                                                                                                                          InterPro; IPR003255; ATP_synt_B_c.
InterPro; IPR000194; ATPase_alpha_beta.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               Arabidopsis thaliana (Mouse-ear cress)
                                                                                                                                                                           EMBL; U96487; AAB64417.1; ALT_INIT.
                                                                                                                                                                                                                                                                                                                                          Query Match 5.6%,
Best Local Similarity 19.1%,
Matches 71; Conservative
   REGULATORY SUBUNIT.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   346 ---YPPGAEKR 353
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   451 ILKYHPKYRKR 461
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               NCBI_TaxID=3702;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   Eukaryota;
                                                                                                                                                                                                                                                                                            SEQUENCE
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  VATB_ARATH
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    8
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    õ
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  q
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                q
```

```
This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation the European Bioinformatics Institute. There are no restrictions on its most by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See http://www.isb-sib.ch/announce/or send an email to license@isb-sib.ch).
                                                                                                                                                                                                               Submitted (JAN-1994) to the EMBL/GenBank/DDBJ databases.
-!-FUNCTION: NONCATALYTIC SUBUNIT OF THE PREXPHERAL VI COMPLEX OF VACUOLAR ATPASE. 18 RESPONSIBLE FOR ACIDIFYING A VARIETY OF INTRACELLULAR COMPARTMENTS IN EUKARYOTIC CELLS.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   119 YNHRLTMHPGQYTQIASPREVVV-----DSAIRDL-AYHDEILSRMKLNEQLNKDAV 169
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   170 LIIHLGGTFEGKKETLDRFRKNYQRLSDSVKARLVLENDDVSWSVQDLLPLCQELNIPLV 229
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   315 TIYERAGRIEGRKGSITQI------PILTMPNDDITHPTPDLTGYITEGQIYID 362
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              230 LDWHHHNIVPGTLREGSLDLMPLIPTIRETWTRKGITQKQH-----YSESADPTAISG 282
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  363 RQLHNRQIYP-----PINVLPSLSRLMKSAIGEGMTRKDHSDVSNQLYANYAIGKDVQA 416
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       255 ITPRIALTTAEYLAYECGKHVLVILTDMSSYADALREVSAAREEVPGRRGYPGYMYTDLA 314
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    158 MNSIARGQKIPLFSAAGLPHNEIAAQICR----QAGL--VKRL--EKTVDLLE----DH 204
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                205 GEDNFAIVFAAMGVNM------ETAQFFKRDFEENGSMERVTLFLNLANDPTIERI 254
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            27 LNTILRSMKERVF-----CSRTCRITTIQRDGLESVKQLGTQNVLDLIKLVEWNH 76
                                                                                                                                                                                                                                                                                                                                                       -!- SUBUNIT: V-ATPASE IS AN HETEROMULTIMERIC ENZYME COMPOSED OF A PERIPHERAL CYPALYTIC VI COMPLEX (MAIN COMPONENTS: SUBUNITS A, B, C, D, E, AND F) ATTACHED TO AN INTEGRAL MEMBRANE VO PROTON PORE COMPLEX (MAIN COMPONENT: THE PROTECLIPID PROTEIN)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          SIMILARITY: BELONGS TO THE ATPASE ALPHA/BETA CHAINS FAMILY.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     5.5%; Score 108.5; DB 1; Length 492; 20.0%; Pred. No. 0.41;
                                                                                                                                           STRAIN=CV. COLUMBIA, TISSUE~Seedling;
Desprez T., Amselem J., Chiapello H., Caboche M., Hofte H.;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     R -> V (IN REF. 2).
2BF630AA79EEAC7C CRC64;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              283 MKRRAHSDRVFDFPPCDPTMDLM-IEAKEKEQAVFELCRRYELQN 326
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               subunit of the vacuolar ATPase from Arabidopsis.";
J. Biol. Chem. 263:17987-17994(1988).
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        77 ---NFGIHFMRVSSDLFPFASHAKYGYTLEFAQSHLEEVG----
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       Hydrolase; ATP synthesis; Hydrogen ion transport.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        Mismatches 123;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                (Rel. 33, Last sequence update) (Rel. 40, Last annotation update)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       Pred. No. 0.41;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       500 AA
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      Pfan, PF00006; ATP-synt_ab, 1.
Pfan, PF01038; ATP_synt_B_c; 1.
PROSITE; PS00152; ATPASE_ALPHA_BETA; 1.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        IPR000194; ATPase_alpha_beta
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      InterPro; IPR003255; ATP_synt_B_c
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            :09
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              (Rel. 33, Created)
(Rel. 33, Last seq
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           492 AA; 54739 MW;
                                                                                                             SEQUENCE OF 241-340 FROM N.A.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     EMBL; J04185; AAC36485.1; -. EMBL; Z29126; CAA82390.1; -.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            Conservative
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       STANDARD;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                PIR; A31886; A31886.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       Best Local Similarity
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     303
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          01-FEB-1996
01-FEB-1996
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      20-AUG-2001
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            :69
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       VATB_CYACA
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            InterPro;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           SEQUENCE
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        Query Match
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         CONFLICT
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                P48413;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     VATB_CYACA
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            Matches
```

us-09-724-296-4.rsp

```
--
--
::
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          Query Match
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          Matches
      qq
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    ōλ
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               qq
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 Οý
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               qq
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           δŏ
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           Op
                                                                                                                                                                                                                                                                                                                                                  between the Swiss Institute of Bioinformatics and thought custation the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See http://www.isb-sib.ch/announce/or send an email to license@isb-sib.ch).
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              12;
                                                                                                                                                                                                                                                                                                                                      This SWISS-PROT entry is copyright. It is produced through a collaboration
                                             Eŭkaryota; Rhodophyta; Bangiophyceae; Porphyridiales; Porphyridiaceae;
                                                                                                                                      "Liegler K., Hauska G., Nelson N.;
"Cyanidium caldarium genes encoding subunits A and B of V-ATPase.";
Biochin. Biochin. Biophys. Acta 1230:202-206(1995).
-i- FUNCTION: NONCATALYIC SUBUNIT OF THE PERIPHERAL VI COMPLEX OF
VACUOLAR ATPASE. V-ATPASE IS RESPONSIBLE FOR ACIDIFYING A VARIETY
OF INTRACELLULAR COMPARTMENTS IN BURARYOTIC CELLS.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       142 PEEMFE-----TGISSIDVMNSIARGOKIPLFSGAGLPHNEVAAQICRQVCLVSTCT 193
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     49 -IQRDGLESVKQLGTQNVLDLIKLVEWNHNFGIHFMRVSSDLFPFASHAKYGYTLEFAQS 107
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   108 HLEEVG------KLANK----YNHRLTMHPGQYTQIASPREVVV-----DSA 144
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               285 LREVSAAREEVPGRRGYPGYMYTDLATIYERAGRVEGRPGSITQL-------PIL 332
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               204 VLENDDVSWSVQDLLPLCQELNIPLVLDWHHHNIVPGTLREGSLDLMPLIPTIRETWTRK 263
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            333 TMPNDDITHPIPDLTGYITEEQIYLDRQLHNRQIYP-----PINVLPSLSRLMKSAIGE 386
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              Gaps
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         5 PREMFDCLDKPIPWRGRLGYACLNTILRSMKERVF-------CSRTCRITT-- 48
                                                                                                                                                                                                                                       SUBUNIT: V-ATPASE IS AN HETEROMULTIMERIC ENZYME COMPOSED OF A PERIPHERAL CATALYTIC VI COMPLEX (MAIN COMPONENTS: SUBUNITS A, B, C, JN F, ATTACHED TO AN INTEGRAL MEMBRANE VO PROTON PORE COMPLEX (MAIN COMPONENT: THE PROTEOLIPID PROTEIN).

SIMILARITY: BELONGS TO THE ATPASE ALPHA/BETA CHAINS FAMILY.
VACUOLAR ATP SYNTHASE SUBUNIT B (EC 3.6.1.34) (V-ATPASE B SUBUNIT);
VACUOLAR PROTON PUMP B SUBUNIT).
Cyanidium caldarium.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     194 LVKRSGKD------ETARFFRQ
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 225 DFEENGAMERVTLFLNLANDPTIERIITPRLALTFAEYLAYEKGKHVLVILTDMSAYADA
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 145 IRDL-AYHDEILSRMKLNEQLNKDAVLIIHLGGTFEGKKETLDRFRKNYQRLSDSVKARL
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         52; Mismatches 129; Indels 106;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             264 GITQKQHYSESADPTAISGMKRRAHSDRVFDFPPCDPTMDLMIEAKEKEQAVF 316
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   387 GMTRKDHSDVSNQLYAAYAMGKDALAMRAV------VGVEALSOEDLLY 429
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            DB 1; Length 500;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                              InterPro; IPR003255; ATPL_synt_B_C.
InterPro; IPR00194; ATPase_alpha_beta.
InterPro; OR001094; ATPase_alpha_beta.
Pfam; PF01038; ATP_synt_B_C; 1.
PR0SITE; PS00152; ATPASE_ALPHA_BETA; 1.
PROSITE; ATP Synthesis; Hydrogen ion transport.
SEQUENCE 500 AA; 56433 MW; 4A882382EEEE9CF5 CRC64;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           Query Match 5.5%; Score 108; DB Best Local Similarity 18.7%; Pred. No. 0.46;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            470 AA.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          PRT;
                                                                                                                             MEDLINE=95345110; PubMed=7619836;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     20-AUG-2001 (Rel. 40, Created)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                 EMBL; U17101; AAA85821.1; -.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         66; Conservative
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          STANDARD;
                                                                                                            SEQUENCE FROM N.A.
                                                                              NCBI_TaxID=2771;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       VATB_ARCFU O29100;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       Matches
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           VATB_ARCFU
ŏλ
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   δ
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                q
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               δy
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              q
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              òγ
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            ò
```

```
This SWISS-PROT entry is copyright. It is produced through a collaboration
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   between the Swiss Institute of Bioinformatics and the EMBL outstation. The European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See http://www.isb-sib.ch/announce/or send an email to license@isb-sib.ch).
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 144 AIRDL-AYHDETLSRMKLNEQLNKDAVLIIHLGGTFEGKKETLDRFRKNYQRLSDSVKAR 202
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               203 LVLENDDVSWSVQDLLIPLCQELNIPLVLDWHHHNIVPGTLREGSLDLMPLIPTIRETWTR 262
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         263 KGITQKQII------YSESAD---PTAISGMKRRAHSDRVFDFPPCDPTMDLMIEAK 309
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    356 EGYTREDHFQWNDQMYAAYAEGVDLRGLVATVGEEALSERDRLF------LKFA 403
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        254 ALREISAAREEVPGRRGYPGYMYTDLATIYERAGRIRGRKGTITQM-------PI 301
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    310 EKEQAVFELCRRYELONPPCPLEI-----MGPEYDQTR-----DGYYPPGAEKRLTAR 357
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           Gaps
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                Hydrolase; ATP synthesis; Hydrogen ion transport; Complete proteome. SEQUENCE 470 AA; 52874 MW; C5782F845F8873E2 CRC64;
20-AuG-2001 (Rel. 40, Last sequence update)
20-AuG-2001 (Rel. 40, Last annotation update)
V-TYPE ATP SYNTHASE BETA CHAIN (EC 3.6.1.34) (V-TYPE ATPASE SUBUNIT
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    Klenk H.-P., Clayton R.A., Tomb J.-F., White O., Nelson K.E., Retchum K.A., Dodson R.J., Gwinn M., Hickey E.K., Peterson J.D., Richardson D.L., Kerlavage A.R., Graham D.E., Kyrpides N.C., Fleischmann R.D., Quackenbush J., Lee N.H., Sutton G.G., Gill S., Kirkness E.F., Dougherty B.A., McKenney K., Adams M.D., Loftus B., Peterson S., Reich C.I., McNeil L.K., Badger J.H., Glodek A., Zhou Overbeek R., Gocayne J.D., Weidman J.F., McDonald L., Utterback T., Cotton M.D., Spriggs T., Artiach P., Kaine B.P., Sykes S.M., Sadow P.W., D'Andrea K.P., Bowman C., Fujil C., Garland S.A., Mason T.M., Olsen G.J., Fraser C.M., Smith H.O., Woese C.R.,
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     "The complete genome sequence of the hyperthermophilic, sulphate-reducing archaeon Archaeoglobus fulgidus.";
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      !- FUNCTION: PRODUCES ATP FROM ADP IN THE PRESENCE OF A PROTON GRADIENT ACROSS THE MEMBRANE. THE ARCHAEAL BETA CHAIN IS A
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  63;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          !- SIMILARITY: BELONGS TO THE ATPASE ALPHA/BETA CHAINS FAMILY.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   5.3%; Score 104.5; DB 1; Length 470;
                                                                                                                                                                                                                                                 Archaea; Euryarchaeota; Archaeoglobales; Archaeoglobaceae;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              93; Indels
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         Pred. No. 0.79;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              45; Mismatches
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 Prosite; PS01038; ATP_Synt_B_c; 1.
PROSITE; PS00152; ATPASE_ALPHA_BETA; 1.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             interpro; IPR003255; ATP_synt_B_c.
Interpro; IPR000194; ATPase_alpha_beta.
                                                                                                                                                                                                                                                                                                                                                                                                                                                 DSM 4304 / ATCC 49558;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  EMBL; AE001023; AAB90073.1; ALT_INIT.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       MEDLINE=98049343; PubMed=9389475;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                Pfam; PF00006; ATP-synt_ab; 1.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    Best Local Similarity 20.2%;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    Nature 390:364-370(1997).
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              51; Conservative
                                                                                                                                                                                                            Archaeoglobus fulgidus.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     REGULATORY SUBUNIT.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  358 KRRSRKEEVEED 369
                                                                                                                                                                                                                                                                                                                                                                                                   SEQUENCE FROM N.A.
                                                                                                                                                                                                                                                                                                                                NCBI_TaxID=2234;
                                                                                                                                                                  ATPB OR AF1167.
                                                                                                                                                                                                                                                                                                                                                                                                                                                     STRAIN=VC-16 /
                                                                                                                                                                                                                                                                                       Archaeoglobus
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    TIGR: AF1167;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        Venter J.C.;
```

S

Qy Db

457 KKEAKPEEATEE 468

g

```
This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation the Buropean Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See http://www.isb-sib.ch/announce/
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       13;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                287 AHSDRVFDFPPCDPTMDLMIEAKEKEQAVFE-LCRRYELQNPPCPLEIMGPEYDQTRDGY 345
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            Thermococcus sp. (strain KI).
Archaea; Euryarchaeota; Thermococcales; Thermococcaceae; Thermococcus.
NCBL_TaxID=35749;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               180 GKKETLDRFRKNYQRLSDSVKARLVLENDDVSWSVQDLLPLCQELNIPLVLDWHHHNIVP 239
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         240 GTLREGSLDLMPLIPTIRETWTRKGITQKQH-----YSES---ADPTAISGMKRR 286
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                139 MNTLVRGQKLPIFSGSGLPHNMLAAQIARQAKVLGËE-------EQFAVVF---- 182
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             SDLFPFASHAKYGYTLE---FAQSHLEEVGKLANK------YNHRLTMHPG 128
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        183 -----AAMGITYEEANFFKKSFEETGAIERAVLFLNLADDPAIERIITPRMALTVA 233
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     129 QYTQIASPREVVV------DSAIRDL-AYHDEILSRMKLNEQLNKDAVLIIHLGGTFE 179
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          / Match 5.3%; Score 104; DB 1; Length 463; Local Similarity 18.9%; Pred. No. 0.85; nes 70; Conservative 55; Mismatches 154; Indels 92; Gaps
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   27 LNTILRSMKERVFCSRTCRITTIQRDGLESVKQLGTQNVLDLIKLVEWNHNFGIHFMRVS 86
                                                                                                                                                                                                                                                                                                                                   Idd T., Kanai S., Inatomi K., Kamagata Y., Maruyama T.;
"Alpha- and beta-subunits of a V-type membrane ATPase in a hyperthermophilic sulfur-dependent archaeum, Thermococcus sp. KI.";
Biochim. Biophys. Acta 1329-13-17(1997).
-!- FUNCTION: PRODUCES ATP FROM ADP IN THE PRESENCE OF A PROTON
GRADIENT ACROSS THE MEMBRANE. THE ARCHAEAL BETA CHAIN IS A
                                                           20-AUG-2001 (Rel. 40, Created)
20-AUG-2001 (Rel. 40, Last sequence update)
20-AUG-2001 (Rel. 40, Last annotation update)
V-TYPE ATP SYNTHASE BETA CHAIN (EC 3.6.1.34) (V-TYPE ATPASE SUBUNIT
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         234 EYLAFDYDMQVLVILTDMTNYAEALREISAAREEVPGRRGYPGYMYTDLATIYERAGRVR
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 342 -----PIDVLPSLSRLMKDGIGKGRTREDHPQLSQQLYAAYAEGRSLRDLVAVVGEEAL
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             -! - SIMILARITY: BELONGS TO THE ATPASE ALPHA/BETA CHAINS FAMILY.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  Hydrolase; ATP synthesis; Hydrogen ion transport.
SEQUENCE 463 AA; 52168 MW; 2F0EC25DA60462DF CRC64;
                  463 AA
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              or send an email to license@isb-sib.ch).
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             PROSITE; PS00152; ATPASE_ALPHA_BETA; 1.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   InterPro; IPR003255; ATP_synt_B_C.
InterPro; IPR000194; ATPARSe_alpha_beta.
Pfam; PF00006; ATP-synt_ab; 1.
Pfam; PF01038; ATP_synt_B_C; 1.
                  PRT;
                                                                                                                                                                                                                                                                                                                 MEDLINE=98036007; PubMed=9370240;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              EMBL; D88772; BAA23343.1; -
                     STANDARD;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                               REGULATORY SUBUNIT.
                                                                                                                                                                                                                                                                                                SEQUENCE FROM N.A.
                     VATB_THESI
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              Query Match
VATB_THESI
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             87
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         Matches
                                                                                                                                                                                 ATPB.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 Op
                                               q
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             δ
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            a
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      ò
```

```
RX ATAIN-BERKBLEY;
RX Adams M.D., Celniker S.E., Holt R.A., Evans C.A., Gocayne J.D.,
RA Adams M.D., Celniker S.E., Holt R.A., Evans C.A., Gocayne J.D.,
RA Adamstides P.G., Scherer S.E., Holt R.A., Cable R.F.,
Sulton G.G., Wortman J.R., Yandell N.D., Zhang O., Chen L.X.,
Brandon R.C., Rogers Y.-H.C., Blazel R.G., Champe M., Pfeiffer B.D.,
RA Abril J.F., Aqbayani A., Baxendale J., Bayarktaroglu L., Beasley E.M.,
Ballew R.M., Bacun A., Baxendale J., Bayarktaroglu L., Beasley E.M.,
Ballew R.M., Bacun M.R., Bouck J., Brokstein P., Brothier P.,
Borkova D., Botchen M.R., Bouck J., Brokstein P., Brothier P.,
Borkova D., Botchen M.R., Bouck J., Brokstein P., Brothier P.,
Borkova D., Botchen M.R., Bouck J., Brokstein P., Brothier P.,
Borkova D., Botcher A., Deng Z., Mays A.D., Dew I., Dietz S.M.,
A Cherry J.M., Cawley S., Dahlke C., Davenport L.B., Davies P.,
Botcher J. Gabriellan A.E., Gary N.S., Gelbart W.M., Classer R.A,
Bodson K., Doup L.E., Downes M., Dugan Rocha S., Punkov B.C., Dunn P.,
RA Glodek A., Gong F., Gornell J.H., Gu Z., Gulbar W.M., Classer R.A,
Alania M., Kalush F., Karpen G. H., Ke Z., Kannison J.A., Ketchum K.A.,
Jalai M., Kalush F., Karpen G. H., Ke Z., Kannison J.A., Ketchum K.A.,
Jalai M., Kalush F., Karpen G. H., Ke Z., Kannison J.A., Ketchum K.A.,
Jalai M., Kalush F., Karpen G. H., Ke Z., Kannison J.A., Ketchum K.A.,
Jalai M., Kalush F., Karpen G. H., Ke Z., Kannison J.A., Ketchum K.A.,
Jalai M., Kalush F., Karpen G. H., Kalush W., Nalson D.L.,
RA Minnel B.E., Kodire C.D., Kraft C., Mortis J., Mornici J., Mornici J., Worlerson D.,
RA Minnel B.E., Kodire C.D., Kraft C., Mortis J., Mornici J., Worlerson D.,
RA Mount S.M., Worden W.A., Li J., Li Z., Liang Y., Lin X.,
Liako P., Lei Y., Levitsky A.A., Li J., Li Z., Liang Y., Lin X.,
Rabson D.R., Worly W., Murphy B., Murphy L., Worley K., Sun E.,
Spirskas R., Moya W., Murphy B., Murphy L., Worley K., Sun E.,
Spirskas R., Woodage T., Worley K.C., Wu D., Yang S., Yan O.,
A Palazzolo M., Pittman G.S., Pan S., Pollard J., Woodage T., Wassarman D.A., Weiss
                                                                                                                                                                                                                       Eukaryota; Metazoa; Arthropoda; Tracheata; Hexapoda; Insecta;
Pterygota; Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;
Ephydroidea; Drosophilidae; Drosophila.
                                                                                                                                                                                                                                                                                                                                                                                                                                                              "Analysis and inactivation of vha55, the gene encoding the vacuolar ATPase B-subunit in Drosophila melanogaster reveals a larval lethal
                                                                                           01-JUL-1993 (Rel. 26, Last sequence update)
COANG-2001 (Rel. 40, Last annotation update)
VACUOLAR ATP SYNTHASE SUBUNIT B (EC 3.6.1.34) (V-ATPASE B SUBUNIT)
(VACUOLAR APPOTON PUMP B SUBUNIT) (V-ATPASE 57 KDA SUBUNIT).
                                                                                                                                                                                                                                                                                                                                                                                                                Davies S.A., Kelly D.C., Goodwin S.F., Yang S., Sozen M.A.,
                       490 AA
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   Biol. Chem. 271:30677-30684(1996).
                                                                                                                                                                                                         Drosophila melanogaster (Fruit fly).
                                                                                                                                                                                                                                                                                                                                                                       STRAIN=EYA; TISSUE=Head;
MEDLINE=97094810; PubMed=8940044;
                     VATB_DROME STANDARD; F
P31409, Q9VG52;
1-JUL-1993 (Rel. 26, Created)
01-JUL-1993 (Rel. 26, Last sequ
                                                                                                                                                                                                                                                                                                                                                                                                                                          Kaiser K., Dow J.A.T.;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 SEQUENCE FROM N.A.
                                                                                                                                                                                                                                                                                                                                                  SEQUENCE FROM N.A.
                                                                                                                                                                                      CG17369
                                                                                                                                                                                                                                                                                                  NCBI_TaxID=7227;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         STRAIN-BERKELEY;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  phenotype.
                                                                                                                                                                                      VHA55 OR
VATB_DROME
```

```
This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation-the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See http://www.isb-sib.ch/announce/or send an email to license@isb-sib.ch).
FUNCTION: NONCATALYTIC SUBUNIT OF THE PERIPHERAL VI COMPLEX OF VACUOLAR ATPASE. V-ATPASE IS RESPONSIBLE FOR ACIDIFYING A VARIETY OF INTERCELLULAR COMPARTMENTS IN EUKARYOTIC CELLS.
SUBUNIT: V-ATPASE IS AN HETREMOULTIMERIC ENZYME COMPOSED OF A PERIPHERAL CATALYTIC VI COMPLEX (MAIN COMPONENTS: SUBUNITS A, B, C, D, E, AND F) ATTACHED TO AN INTEGRAL MEMBRANE VO PROTON PORE COMPLEX (MAIN COMPONENT: THE PROTEOLIPID PROTEIN).
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 128 DFLDIQGQPINPWSRIYPEEMIQTGISAIDVMNSIARGQKIPIFSAAGLPHNEIAAQICR 187
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      46 ITTIQRDGLESVKQLGTQNVLDLIKLVEWNHNFGIHFMRVSSDLFPFASHAKYGYTLEFA 105
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             106 QSHLEEVGKLANK--------YNHRLTMHPGQYTQIASPREVVV-----D 142
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      143 SAIRDL-AYHDEILSRMKLNEQLNKDAVLIIHLGGTFEGKKETLDRFRKNYQRLSDSVKA 201
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 285 EALREVSAAREEVPGRRGFPGYMYTDLATIYERAGRVEGRNGSITQI------P 332
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          333 ILTMPNDDITHPIPDLTGYITEGOIYVDRQLHNRQIYP-----PVNVLPSLSRLMKSAI 386
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          225 KODFEENGSMENVCLFLNLANDPTIERIITPRLALTAAEFLAYOCEKHVLVILTDMSSYA 284
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 202 RLVLENDDVSWSVQDLLPLCQELNIPLVLDWHHHNIVPGTLREGSLDLMPLIPTIRETWT 261
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         86; Gaps
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                7 EMFDCLDKPI-PW-----CSRLGYACLNTILRSMKERVF------CSRTCR 45
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             GOG4_HUMAN STANDARD; PRT; 2230 AA. 01436; 013270; 013654; 22-40436; 013270; 013654; 20-40G-2001 (Rel. 40, created) 20-AUG-2001 (Rel. 40, Last sequence update) 20-AUG-2001 (Rel. 40, Last annotation update) 60LGI AUTOANTIGEN, GOLGIN SUBPAMILY A 4 (TRANS-GOLGI P230) (256 KDA GOLGIN) (GOLGIN-245) (72.1 PROTEIN).
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     Homo sapiens (Human).

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;

Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
...
                                                                                                                                                                 SIMILARITY: BELONGS TO THE ATPASE ALPHA/BETA CHAINS FAMILY.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            DB 1; Length 490;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       48; Mismatches 114; Indels
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              Hydrolase; ATP synthesis; Hydrogen ion transport.
SEQUENCE 490 AA; 54549 MW; 89651BE1B5A53618 CRC64;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         Query Match 5.3%; Score 104; DB Best Local Similarity 19.7%; Pred. No. 0.92;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           Pfam; PF00006; ATP-synt_ab; 1.
Pfam; PF01038; ATP-synt_Bc; 1.
PROSITE; PS00152; ATPASE_ALPHA_BETA; 1.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                               InterPro; IPR001255; ATP_synt_B_c.
InterPro; IPR000194; ATPase_alpha_beta.
                                                                                                                                                                                                                                                                                                                                                                                                              EMBL; AE003696; AAF54837.1; -. PIR; S25167; S25167. FlyBase; FBgn0005671; Vha55.
                                                                                                                                                                                                                                                                                                                                                                                           EMBL; X67839; CAA48034.1; -.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       61; Conservative
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     262 RKGITQKQH 270
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                387 GEGMTRKDH 395
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         Query Match
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              GOLGA4.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    RESULT 9
GOG4_HUMAN
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       Matches
       ŏ
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             δ
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   οy
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               q
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               ŏ
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              Dp
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   δý
```

```
This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation - the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See http://www.isb-sib.ch/announce/or send an email to license@isb-sib.ch).
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         17;
                                                                                                                                                                                                                                                                                                                                                                                                                   -!- ALTERNATIVE PRODUCIS: DIFFERENTS ISOFORMS MAY ARISE BY ALTERNATIVE
                                                                                                                                                                                                                                                                                                                                                                                                                                              -!- DISEASE: ANTIGEN IN THE AUTOIMMUNE DISEASE SJOEGREN'S SYNDROME AND
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            TPYKGGNLYHTDVSLFGEPTEFEYLRKVLFEY -> HLTKV
                                                                                                                                                                                                                                                                                                                                                                                       1- SUBCELLULAR LOCATION: ASSOCIATED WITH THE CYTOPLASMIC FACE OF THE
                                                                                                                                                                                                                                                                                                                                  Balague C.;
Thesis (1994), Instituto municipal de investigacion medica, Spain.
-!- FUNCTION: MAY PLAY A ROLE IN VESICULAR TRANSPORT FROM THE TRANS-
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          AICTIRMSHSLENLPNLSICEKCFLSI (IN ISOFORM
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         Gaps
                         MEDLINE=96215236; PubMed=8626529;
Erlich R., Gleeson P.A., Campbell P., Dietzsch E., Toh B.-H.;
"Molecular characterization of trans-Golgi p230: a human peripheral
membrane protein encoded by a gene on chromosome 6p12-22 contains
extensive coiled-coil alpha helical domains and a granin motif.";
J Blol. Chem. 271:8328-8337(1996).
                                                                                                                                                                                                                           Fritzler M.J., Lung C.-C., Hamel J.C., Griffith K.J., Chan E.K.L., Molecular characterization of golgin-245, a novel Golgi complex protein containing a granin signature."; J. Biol. Chem. 270:31262-31268(1995).
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  SWLRSSS (IN ISOFORM 4).
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        152; Indels 101;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             5.2%; Score 103.5; DB 1; Length 2230;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           coil; Alternative splicing.
                                                                                                                                                         Submitted (NOV-1994) to the EMBL/GenBank/DDBJ databases
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   COLLED COIL (POTENTIAL).
COLLED COIL (POTENTIAL).
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   POTENTIAL)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     MISSING (IN ISOFORM 2).
MISSING (IN ISOFORM 3).
FTSPRSGIF -> SWLRSSS (I
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                COILED COIL (POTENTIAL)
COILED COIL (POTENTIAL)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          ; Pred. No. 8.2; 71; Mismatches
                                                                                                                                                                                                                 MEDLINE=96125.112; PubMed=8537393;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  261139 MW;
                                                                                                                                                                                  SEQUENCE OF 131-2230 FROM N.A.
                                                                                                                                                                                                                                                                                                   SEQUENCE OF 524-672 FROM N.A.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    EMBL; U41740; AAC50434.1; -. EMBL; X82834; CAA58041.1; -.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               EMBL; U31906; AAC51791.1; -...
EMBL; X76942; CAA54261.1; -...
MIM; 602509; -...
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       stack; Antigen; Coiled
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          19.4%;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          InterPro; IPR000237; GRIP. Pfam; PF01465; GRIP; 1.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      Best Local Similarity 19.4 Matches 78; Conservative
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    237
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                121.4
2152
2185
                                                                                                                                                                                                                                                                                                                TlSSUE=Gastric fundus;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              188
220
276
276
584
628
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  682
                                                                                                                                                                                                                                                                                                                                                                                                                                                             IN HEPATITIS B.
             SEQUENCE FROM N.A.
                                                                                                                            SEQUENCE FROM N.A.
                                                                                                                                                                                                                                                                                                                                                                                                     GOLGI MEMBRANE
                                                                                                                                                                                                     TISSUE=Placenta;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              188
220
220
276
584
628
630
682
2230
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            1259
2154
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 1033
                                                                                                                                                                                                                                                                                                                                                                                                                                  SPLICING.
                                                                                                                                           Seelig H.P.;
                                                                                                                                                                                                                                                                                                                                 Balaque C
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  VARSPLIC
VARSPLIC
VARSPLIC
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             DOMAIN
VARSPLIC
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             Query Match
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                CONFLICT
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             CONFLICT
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             CONFLICT
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          CONFLICT
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     CONFLICT
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    DOMAIN
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   DOMAIN
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       Golgi
```

```
This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See http://www.isb-sib.ch/announce/or send an email to license@isb-sib.ch).
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       h. Biochem. Biophys. 291:92-99(1991).
FUNCTION: NONCATALYTIC SUBUNIT OF THE PERIPHERAL VI COMPLEX OF FUNCTION: NONCATALYTIC SUBUNIT OF THE PERIPHERAL VI COMPLEX OF YACUOLAR ATPASE. V-ATPASE IS RESPONSIBLE FOR ACIDIFYING A VARIETY OF INTRACELULIAR COMPARTMENTS IN BUKARYOTIC CELLS.
SUBUNIT: V-ATPASE IS AN HETEROMULTIMERIC ENZYME COMPOSED OF A PERIPHERAL CATALYTIC VI COMPLEX (MAIN COMPONENTS: SUBUNITS A, B, C, D, E, AND F) ATTACHED TO AN INTEGRAL MEMBRANE VO PROTON PORE COMPLEX (MAIN COMPONENT: THE PROTEDLIPLD PROTEIN).
SIMILARITY: BELONGS TO THE ATPASE ALPHA/BETA CHAINS FAMILY.
                                                                                                                                                                                                                 1852 LEQKIKELDSCLVRQKEVHRVEMEELTSKYEKLQALQQMDGRNKPTELLEENTEEKSKSH 1911
                                                                                                                                      1682 QEREREVHILEEKLKSVESSQSETLIVPRSAKNVAAYT---EQEEADSQGCVQKTYEEKI 1738
                                                                                                                                                                                                                                                                                                              1912 LVQPKLLSNMEAQ-HNDLEFKLAGAE------REKQKLGKEIVRLQKDLRMLRKEHQ 1961
                                              1622 LKALEDRLESESAAKLAELKRKAEQKIAAIKKQLLSQMEEKEEQYKKGTESHLSELNTKL 1681
                                                                                                                                                                                    124 TMHPGQYTQ-----IASPREVVVDSAIRDLAYHDEILSRMKLNE-QLNKDAVLIIHLG 175
                                                                                                                                                                                                                                                                                                                                                                                                                                                                       325
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      Eukaryota; Metazoa; Arthropoda; Trachata "mem";
Pterygota; Neoptera; Endopterygota; Lepidoptera; Glossata; Ditrysia;
Noctuoidea; Noctuidae; Heliothinae; Heliothis.
----GTONVLDLIKLVE 73
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         01-JUL-1993 (Rel. 26, Last sequence update)
20-AUG-2001 (Rel. 40, Last annotation update)
VACUOLAR ATP SYNTHASE SUBUNIT B (EC 3.6.1.34) (V-ATPASE B SUBUNIT)
                                                                                                                                                                                                                                                                                                                                                                                223 -----ELNIPLVLDWHHHNIVPGTLR----EGSLDLMPLIPTIRETWTRKGITQKQHYSE
                                                                                                                                                                                                                                                                                                                                                                                                                                                                       274 SADPTAISGMKRRAHSDRVFDFPPCDPTMDLMIEAKEKEQAVFELCR-----RYELQ
                                                                                          WNHNFGIHFMR----VSSD-----LFPFASHAKYGYTLEFAQSHLEEVGKLANKYNHRL
                                                                                                                                                                                                                                                                                   176 GTFEGKKETLDRFRKNYQRL-----SDSVKARLVLEN--DDVSWSVQDLLPLCQ-
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          "Molecular cloning and characterization of the B subunit of a vacuolar H(+)-ATPase from the midgut and Malpighian tubules of
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  ----EKIKOEOEDLE 1987
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    326 NPPCPLEIMGPEYDQTRDGYYPPGAEKRLTARKRRSRKEEVE 367
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            Heliothis virescens (Noctuid moth) (Owlet moth).
  31 LRSMKERVFCSRTCRITTIQRDGLESV----KQL----
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     494 AA
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     PRT;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              MEDLINE=92027804; PubMed=1834020;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           (VACUOLAR PROTON PUMP B SUBUNIT).
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   01-JUL-1993 (Rel. 26, Created)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     EMBL; S61797; AAB20098.1; -. PIR; S18395; S18395.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                1962 Q---ELEILKKEYDOERE----
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     STANDARD;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             Helicoverpa virescens.";
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           3ill S.S., Ross L.S.;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    SEQUENCE FROM N.A.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       VATB_HELVI
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        Arch.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 VATB_HELVI
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             g
                                                                                                                                                                                                                                     qq
                                                                                                                                                                                                                                                                                                                                 a
                                                                                                                                                                                                                                                                                                                                                                                ò
                                                                                                                                                                                                                                                                                                                                                                                                                            g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                            ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       ò
                                                                                                                                          g
                                                                                                                                                                                                                                                                                     ò
                                                                                                 á
```

```
13;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         "Primary structure of V-ATPase subunit B from Manduca sexta midgut."; Biochim. Biophys. Acta 1132:67-71(1992).
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             ----QAGL--VKIPG-KSVLD-----DHEDNFAIVFAAMGVNM------ETARFF 228
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          QSHLEEVGKLANK------D 142
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   229 KQDFEENGSMENVCLFLNLANDPTIERIITPRLALTAAEFLAYQCEKHVLVILTDMSSYA 288
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              289 EALREVSAAREEVPGRRGFPGYMYTDLATIYERAGRVEGRNGSITQI------P 336
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           202 RLVLENDDVSWSVQDLLPLCQELNIPLVLDWHHHNIVPGTLREGSLDLMPLIPTIRETWT 261
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            337 ILTMPNDDITHPIPDLTGYITEGQIYVDRQLHNRQIYP-----PVNVLPSLSRLMKSAI 390
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            46 ITTIQRDGLESVKQLGTQNVLDLIKLVEWNHNFGIHFMRVSSDLFPFASHAKYGYTLEFA 105
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         143 SAIRDL-AYHDEILSRMKLNEQLNKDAVLIIHLGGTFEGKKETLDRFRKNYQRLSDSVKA 201
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    OF INTRACELLULAR COMPARTMENTS IN EUKARYOTIC CELLS. THE B SUBUNIT IS NONCATALYTIC BUT COMBINES WITH OTHER SUBUNITS TO FORM THE CATALYTIC COMPLEX. V-ATPASE IS RESPONSIBLE FOR ENERGIZING
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   132 DFLDIQGQPINPWSRIYPEEMIQTGISAIDVMNSIARGQKIPIFSAAGLPHNEIAAQICR 191
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        ELECTROPHORETIC K+/2H+ ANTIPORT BY GENERATING A TRANSMEMBRANE VOLTAGE OF MORE THAN 200MV.

-!- SUBUNIT: V-ATPASE IS AN HETEROMULTIMERIC ENZYME COMPOSED OF PERIPHERAL CATACHTIC V. COMPLEX (MAIN COMPONENTS: SUBUNITS A, B, C, D, E, AND F) ATTACHED TO AN INTEGRAL MEMBRANE VO PROTON PORE COMPLEX (MAIN COMPONENT: THE PROTECUIPID PROTEIN).

-!- SIMILARITY: BELONGS TO THE ATPASE ALPHA/BETA CHAINS FAMILY.
                                                                                                                                                                                                                                                                                                                                                                                                    86; Gaps
                                                                                                                                                                                                                                                                                                                                                                                                                                                                         7 EMFDCLDKPI-PW-----CSRTCR 45
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             Pterygota; Neoptera; Endopterygota; Lepidoptera; Glossata; Ditrysia;
Sphingiodea; Sphingidae; Sphinginae; Manduca.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   Novak F.J., Graf R., Waring R.B., Wolfersberger M.G., Wieczorek H.,
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   Last sequence update)
Last annotation update)
SUBUNIT B (EC 3.6.1.34) (V-ATPASE B SUBUNIT)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    Manduca sexta (Tobacco hawkmoth) (Tobacco hornworm).
Eukaryota; Metazoa; Arthropoda; Tracheata; Hexapoda; Insecta
                                                                                                                                                                                                                                                                                                                               DB 1; Length 494;
                                                                                                                                                                                                                                                                                                                                                                                                        Indels
                                                                                                                                                                                 Hydrolase; ATP synthesis; Hydrogen ion transport.
SEOUENCE 494 AA; 54893 MW; 2D20C581D8AE4C76 CRC64;
                                                                                                                                                                                                                                                                                                                                                                                                        Mismatches 114;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           494 AA.
                                                                                                                                                                                                                                                                                                                               5.2%; Score 103; DB 19.7%; Pred. No. 1.1;
                                                                 Pfam; PF00006; ATP-synt_ab; 1.
Pfam; PF01038; ATP_synt_B_c; 1.
PROSITE; PS00152; ATPASE_ALPHA_BETA; 1.
                                  IPR000194; ATPase_alpha_beta.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           PRT;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          TISSUE=Midgut;
MEDLINE=92379093; PubMed=1387326;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               (VACUOLAR PROTON PUMP B SUBUNIT).
InterPro; IPR003255; ATP_synt_B
                                                                                                                                                                                                                                                                                                                                                                                                        48;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                01-JUL-1993 (Rel. 26, Created)
01-JUL-1993 (Rel. 26, Last sequ
20-AUG-2001 (Rel. 40, Last anno
                                                                                                                                                                                                                                                                                                                                                                                                        Conservative
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               STANDARD;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        VACUOLAR ATP SYNTHASE
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               SEQUENCE FROM N.A.
                                                                                                                                                                                                                                                                                                                                                                     Similarity
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 262 RKGITQKQH 270
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          391 GEGMTRKDH 399
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      NCBI_Tax1D=7130;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  Harvey W.R.;
                                                                                                                                                                                                                                                                                                                                                                                                    61;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               VATB_MANSE
                                      InterPro;
                                                                                                                                                                                                                                                                                                                                   Query Match
                                                                                                                                                                                                                       SEQUENCE
                                                                                                                                                                                                                                                                                                                                                                         Best Local
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            VATB_MANSE
                                                                                                                                                                                                                                                                                                                                                                                                        Matches
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               The property of the property o
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   δò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             qq
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      δŏ
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       qq
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                δ
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         QQ
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      δλ
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        qq
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     Ο̈́λ
```

```
This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation - the Buropean Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See http://www.isb-sib.ch/announce/or send an email to license@isb-sib.ch).
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    13;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    192 -----QAGL--VKIPG-KSVLD-----DHEDNFAIVFAAMGVNM------ETARFF 228
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             -----YNHRLTMHPGQYTQIASPREVVV-----D 142
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      132 DFLDIQGOPINPWSRIYPEEMIQTGISAIDVMNSIARGOKIPIFSAAGLPHNEIAAQICR 191
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   46 ITTIQRDGLESVKQLGTQNVLDLIKLVEWNHNFGIHFMRVSSDLFPFASHAKYGYTLEFA 105
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          229 KODFEENGSMENVCLFLNLANDPIIERIITPRLALTAAEFLAYQCEKHVLVILTDMSSYA 288
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                289 EALREVSAAREEVPGRRGFPGYMYTDLATIYERAGRVEGRNGSITQI------P 336
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              SAIRDL-AYHDEILSRMKLNEQLNKDAVLIIHLGGTFEGKKETLDRFRKNYQRLSDSVKA: 201
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         202 RLVLENDDVSWSVQDLLPLCQELNIPLVLDWHHHNIVPGTLREGSLDLMPLIPTIRETWT 261
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         7 EMFDCLDKPI-PW------RGRLGYACLNTILRSMKERVF------CSRTCR 45
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               "Selectively amplified expression of an isoform of the vacuolar H(+)-ATPase 56-Kilodalton subunit in renal intercalated cells.";
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    Gaps
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    01-JUL-1993 (Rel. 26, Created)
01-JUL-1993 (Rel. 26, Last sequence update)
20-AUG-2001 (Rel. 40, Last annotation update)
VACUCLAR ATP SYNTHASE SUBBNIT B, KIDNEY ISOFORM (EC 3.6.1.34) (V-ATPASE B1 SUBUNIT) (VACUOLAR PROTON PUMP B ISOFORM 1) (ENDOMEMBRANE PROTON PUMP 58 KDA SUBUNIT).
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Buteleostomi; Mammalia; Eutheria; Cetartiodactyla; Ruminantia; Pecora; Bovoidea;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                86;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   DB 1; Length 494;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             MEDLINE-92228822; PubMed-1373501;
Nelson R.D., Guo X.-L., Masood K., Brown D., Kalkbrenner M.,
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       Ouery Match 5.2%; Score 103; DB 1; Length 49
Best Local Similarity 19.7%; Pred. No. 1.1;
Matches 61; Conservative 48; Mismatches 114; Indels
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               Pfam; PF00006; ATP-synt_ab; 1.
Pfam; PF01038; ATP-synt_ab; 1.
PR051TE; PF00122; ATPASE_ALPHA BETA; 1.
HYDROLASE, ATP Synthesis; Hydrogen ion transport.
SEQUENCE 494 AA; 54904 MW; 8A8A3A2C04A9CFD6 CRC64;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                513 AA.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        SEQUENCE FROM N.A., AND SEQUENCE OF 504-513.
                                                                                                                                                                                                                                                                                                                                                                                                                 InterPro; IPR003255; ATP_synt_B_c.
InterPro; IPR000194; ATPase_alpha_beta.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                PRT;
                                                                                                                                                                                                                                                                                                                EMBL; X64354; CAA45706.1; -.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     106 QSHLEEVGKLANK------
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        STANDARD;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  Bovidae; Bovinae; Bos.
                                                                                                                                                                                                                                                                                                                                                         PIR; S19206; S19206.
PIR; S24387; S24387.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             Bos taurus (Bovine).
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                262 RKGITQKQH 270
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         391 GEGMTRKDH 399
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               NCBI_TaxID=9913;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                TISSUE=Kidney;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    VAB1_BOVIN
P31407;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              Gluck S.;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                ATP6B1
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   VAB1_BOVIN
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                ř,
        SOUR DE PRESENTATION DE PRESEN
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      a
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    qq
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          Q
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               δ
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 δ
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        qq
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         ŏ
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              qq
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            ŏλ
```

```
This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See http://www.isb-sib.ch/announce/or send an email to license@isb-sib.ch).
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      12;
                                                             VACUGLAR ATPASE. V-ATPASE IS RESPONSIBLE FOR ACIDIFYING A VARIETY OF INTRACELLULAR COMPARTMENTS IN EUKARYOTIC CELLS.
-!- SUBUNIT: V-ATPASE IS AN HETEROMULTIMERIC ENXYME COMPOSED OF A PRIPHERAL CATALYTIC VI COMPLEX (MAIN COMPONENTS: SUBUNITS A, B, C, D, E, AND F) ATTACHED TO AN INTEGRAL MEMBRANE VO PROTON PORE COMPLEX (MAIN COMPONENT: THE PROTECLIPID PROTEIN).
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    151 HDEILSRMKLNEQLNKDAVLIIHLGGTFEGKKETLDRFRKNYQRLSDSVKARLVLENDDV 210
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             160 PEEMIETGISPID-----QMNSIARGQKIPIFSAAGLPHNEIAAQICR----QAGL 206
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      55 ESVKQLGTQNVLDLIKLVEWNHNFGIHFMRVSSDLFPFASHAKYGYTLEFAQSHLEEVGK 114
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                207 --VKK--SKAVLDY----HDDNFALVFAAMGVNM------ETARFFKSDFEQNGT 247
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             115 LANK------DSAIRDL-AYNHRLTMHPGQYTQIASPREVVV------DSAIRDL-AY 150
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       248 MGNVCLFLNLANDPTIERIITPRLALTTAEFLAYQCEKHVLVILTDMSSYAEALREVSAA 307
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              308 REEVPGRRCFPGYMYTDLATIYERAGRVEGRGGSITQI-------PILTMPNDDI 355
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            211 SWSVQDLLPLCQELNIPLVLDWHHHNIVPGTLREGSLDLMPLIPTIRETWTRKGITQKQH 270
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 356 THPIPDLTGFITEGOIYVDRQLHNRQIYP-----PINVLPSLSRLMKSAIGEGMTRKDH 409
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  84; Gaps
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Buteleostomi;
Archosauria; Aves; Neognathae; Galliformes; Phasianidae; Phasianinae;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   5 PREMFDCLDKPIPWRGRLGYACLNTILRSMKERVF------CSRTCRITTIQRDGL 54
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             Hydrolase; ATP synthesis; Hydrogen ion transport; Multigene family. SEQUENCE 513 AA; 56746 MW; 4C6DBE4E12570A25 CRC64;
Proc. Natl. Acad. Sci. U.S.A. 89:3541-3545(1992).
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              01-0CT-1996 (Rel. 34, Created)
01-0CT-1996 (Rel. 34, Last sequence update)
20-AUG-2001 (Rel. 40, Last annotation update)
VACUOLAR APP SYMPHIASE SUBUNIT B (EC 3.6.1.34) (V-ATPASE B SUBUNIT)
(VACUOLAR APP SYMPHIASE SUBUNIT B.
                                                                                                                                                                                                                                                                                                                                                                               -!- TISSUE SPECIFICITY: KIDNEY CORTEX AND MEDULLA.
-!- SIMILARITY: BELONGS TO THE ATPASE ALPHA/BETA CHAINS FAMILY.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            5.2%; Score 102; DB 1; Length 513; 19.7%; Pred. No. 1.4;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  49; Mismatches 108; Indels
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         453 AA.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         InterPro; IPR003255; ATP_synt_B_c.
InterPro; IPR000194; ATPase_alpha_beta.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   Pfam: PF00006; ATP-synt_ab; 1.
Pfam: PF01038; ATP-synt_B_c; 1.
PROSITE: PS00152; ATPASE_ALPHA_BETA; 1.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         PRT;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 Pred.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            EMBL; M88691; AAA30394.1; -.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        Query Match
Best Local Similarity 19.78
***Irhas 59; Conservative
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     STANDARD;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               Gallus gallus (Chicken).
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                PIR; C44138; C44138.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              NCBI_Tax1D=9031;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     VATB_CHICK
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      P49712;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        VATB_CHICK
NO NEW WAY OF COURSE OF CO
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       Óζ
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      δò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                a
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               qq
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     {}_{\rm QY}
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    οy
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              q_{0}
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            Οy
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 q
```

SECUENCE FROM N.A.

```
NCBI_TaxID=29292;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        VATB_CAEEL
                                                         Heilig R.;
                                                                                                                                                                                                                                                                                                                                InterPro;
                                                                                                                                                                                                                                                                                                                                                                                                      SEQUENCE
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       019626;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              F20B6.2
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            VATB_CAEEL
                                                                                                                                                                                                                                                                                                                                                                                                                                                                          Matches
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            11D
AC
DT
DT
DE
CO
OC
  qq
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              q
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    qq
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             Öλ
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          a
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             δλ
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  Οý
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          ò
                                                                                                                                                                                               This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation-the Buropean Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See http://www.isb-sib.ch/announce/
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         14;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      142 ----- DSAIRDL-AYHDEILSRMKLNEQLNKDAVLIIHLGGTFEGKKETLDRFRKNY 192
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            193 QRLSDSVKARLVLENDDVSWSVQDLLPLCQELNIPLVLDWHHHNIVPGTLREGSLDLMPL 252
                                                                                   VACUOLAR ATPASE. V-ATPASE IS RESPONSIBLE FOR ACIDIFYING A VARIETY OF INTRACELLULAR COMPARTMENTS IN BUKARYOTIC CELLS.
SUBUNIT: V-ATPASE IS AN HETEROMULTHERIC ENZYME COMPOSED OF A PERIPHERAL CATALYTIC VI COMPLEX (MAIN COMPONENTS: SUBUNITS A, B, C, D, E, AND F) ATTACHED TO AN INTEGRAL MEMBRANE VO PROTON PORE COMPLEX (MAIN COMPONENT: THE PROTEOLIPID PROTEIN).
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     141 NEIAAQICR-----QAGL--VKK--SKDVMDYSE-----ENFAIVFAAMGVNM----- 179
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   97 KYGYTLEFAQSHLEEVGKLANK------YNHRLTMHPGQYTQIASPREVVV 141
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          293 ------PILTMPNDDITHPIPDLTGYITEGQIYVDRQLHNRQIYP-----PINVLPS 338
                            Gaps
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    2 DHAP----REMFDCLDKPIPWRGRL------GYAC---LNTILRSMKERVF----- 39
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           Pyrococcus abyssi.
Archaea; Euryarchaeota; Thermococcales; Thermococcaceae; Pyrococcus.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                20-AUG-2001 (Rel. 40, Last sequence update)
20-AUG-2001 (Rel. 40, Last annotation update)
V-TYPE ATP SYNTHASE BETA CHAIN (EC 3.6.1.34) (V-TYPE ATPASE SUBUNIT
                                                                                                                                                                                                                                                                                                                                                                                                      Hydrolase, ATP synthesis; Hydrogen ion transport; Multigene family.
SEQUENCE 453 AA; 50225 MW; 463B85F93DFFC962 CRC64;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            40 ---CSRTCRITTIQRDGLESVKQLGTQNVLDLIKLVEWNHNFGIHFMRVSSDLFPFASHA
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               180 ---ETARFFKSDFEENGSMDNVCLFLNLANDPTIERIITPRLALTTAEFLAYQCEKHVLV
                                                                                                                                                                      SIMILARITY: BELONGS TO THE ATPASE ALPHA/BETA CHAINS FAMILY.
                                                                                                                                                                                                                                                                                                                                                                                                                                                              Length 453;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         Indels
                                                                                                                                                                                                                                                                                                                                                                                                                                                              DB 1;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        Mismatches 111;
                                                                                                                                                                                                                                                                                                                                                                                                                                                             n 5.1%; Score 101.5; D
Similarity 19.2%; Pred. No. 1.3;
                                                                                                                                                                                                                                                                                                                                                         Pfam; PF00006; ATP-synt_ab; 1.
Pfam; PF01038; ATP_synt_B_c; 1.
PROSITE; PS00152; ATPASE_ALPHA_BETA; 1.
                                                                                                                                                                                                                                                                                       or send an email to license@isb-sib.ch).
                                                                                                                                                                                                                                                                                                                             InterPro; IPR003255; ATP_synt_B_c.
InterPro; IPR000194; ATPase_alpha_beta.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       PRT;
  STRAIN=WHITE LEGHORN; TISSUE=Bone;
MEDLINE=95369682; PubMed=7642089;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         55;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     20-AUG-2001 (Rel. 40, Created)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        253 IPTIRETWTRKGITQKQH 270
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   339 LSRLMKSAIGEGMTRKDH 356
                                                                                                                                                                                                                                                                                                                   EMBL; U20766; AAA82983.1;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           Conservative
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          STANDARD;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            ATPB OR PABI186.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         61;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        VATB_PYRAB
Q9UXU8;
                                                                                                                                                                                                                                                                                                                                                                                                                                                             Query Match
                                                                                                                                                                                                                                                                                                                                                                                                                                                                            Best Local
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              RESULT 14
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            VATB_PYRAB
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         Matches
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          õ
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            qq
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     q
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               ŏ
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            q
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                g
  a
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   ò
```

```
This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (see http://www.isb-sib.ch/announce/or send an email to license@isb-sib.ch).
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     12;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      ----YNHRLTMHPG 128
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 185 ------AAMGTTYEEANFFKKSFEETGAIERAVLFLNLADDPAIERIITPRMALTVA 235
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          129 QYTQIASPREVVV------DSAIRDL-AYHDEILSRMKLNEQLNKDAVLIIHLGGTFE 179
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   236 EYLAFDYDMQVLVILTDMTNYCEALREISAAREEVPGRRGYPGYMYTDLATIYERAGRIR 295
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    180 GKKETLDRFRKNYQRLSDSVKARLVLENDDVSWSVQDLLPLCQELNIPLVLDWHHHNIVP 239
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      240 GTLREGSLDLMPLIPTIRETWTRKGTTQKQH------YSES---ADPTAISGMKRR 286
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           344 -----PIDVLPSLSRLMKDGIGKGRTREDHPQLAQQLYAAYAEGKSLRDLVAVVGEEAL 397
                                                                                                       "Pyrococcus abyssi genome sequence: insights into archaeal chromosome
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   141 MNTLVRGQKLPIFSG-----SGLPH-NQLAAQ-IARQAKVLGEEESFAVVF---- 184
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         Gaps
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          27 LINTILRSMKERVFCSRTCRITTIQRDGLESVKQLGTQNVLDLIKLVEWNHNFGIHFMRVS 86
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        Eukaryota; Metazoa; Nematoda; Chromadorea; Rhabditida; Rhabditoidea;
Rhabditidae; Peloderinae; Caenorhabditis.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          Hydrolase; ATP synthesis; Hydrogen ion transport; Complete proteome.
SEQUENCE 465 AA: 52224 MW; 873F500BD4C5EDFD CRC64;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  B
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             01-NOV-1997 (Rel. 35, Last sequence update)
20-AUG-2001 (Rel. 40, Last annotation update)
PROBABLE VACUOLAR ATP SYNTHASE SUBUNIT B (EC 3.6.1.34) (V-ATPASE SUBUNIT) (VACUOLAR PROTON PUMP B SUBUNIT).
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             83;
                                                                                                                                                                                                                   FUNCTION: PRODUCES ATP FROM ADP IN THE PRESENCE OF A PROTON GRADIENT ACROSS THE MEMBRANE. THE ARCHAEAL BETA CHAIN IS A
                                                                                                                                                                                                                                                                                        REGULATORY SUBUNIT.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        Length 465;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     46; Mismatches 119; Indels
                                                                                                                                                 structure and evolution.";
Submitted (JUL-1999) to the EMBL/GenBank/DDBJ databases.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            DB 1;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               87 SDLFPFASHAKYGYTLE---FAQSHLEEVGKLANK----
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        Ouery Match 5.1%; Score 101.5; E
Best Local Similarity 19.2%; Pred. No. 1.3;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 491 AA
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                IPR000194; ATPase_alpha_beta.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     Pfam; PF01038; ATP_synt_B_c; 1.
PROSITE; PS00152; ATPASE_ALPHA_BETA; 1.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   EMBL; AJ248288; CAB50665.1; -.
InterPro; IPR003255; ATP_synt_B_
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          01-NOV-1997 (Rel. 35, Created)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      Pfam; PF00006; ATP-synt_ab;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 59; Conservative
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 STANDARD;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     Caenorhabditis elegans.
SEQUENCE FROM N.A.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        287 AHSDRVF 293
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             398 SETDKKY 404
                                      STRAIN=ORSAY;
```

NCBI_TaxID=6239;

```
This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation-the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial
                                                                                                                                                                                                                                                                                                                                                                                                                                              entities requires a license agreement (See http://www.isb-sib.ch/announce/or send an email to license@lsb-sib.ch).
                                                                                                          Submitted (DEC-1995) to the EMBL/GenBank/DDBJ databases.
-!- FUNCTION: NONCATALYTIC SUBUNIT OF THE PERIPHERAL VI COMPLEX OF VACUOLAR ATPASE. V-ATPASE IS RESPONSIBLE FOR ACIDIFYING A VARIETY OF INTRACELLULAR COMPARTMENTS IN BUKARYOTIC CELLS (BY SIMILARITY).
-!- SUBUNIT: V-ATPASE IS AN HETEROMUJINERIC ENZYME COMPOSED OF A PERIPHERAL CATALYTIC VI COMPLEX (MAIN COMPONENT: SUBUNITS A, B, C, D, E, AND F) ATTACHED TO AN INTEGRAL MEMBRANE VO PROTON PORE COMPLEX (MAIN COMPONENT: THE PROTEDLIPIN (BY SIMILARITY).
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      131 DFLDINGOPINPWSRIYPEEMIQTGISAIDVMNSIARGOKIPIFSA-----SGLP 180
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    181 HNEIAAQIVRQG--GLVQLPDRPHEQTNFAIVFAAMGVNM------ETARFFKQDF 228
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    110 EEVGKLANK------DSAIR 146
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  147 DL-AYHDEILSRMKLNEQLNKDAVLIIHLGGTFEGKKETLDRFRKNYQRLSDSVKARLVL 205
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   289 EVSAAREEVPGRRGFPGYMYTDLATIYERAGRVEGRDGSITQI-------PILTM 336
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 206 ENDDVSWSVQDLLPLCQELNIPLVLDWHHHNIVPGTLREGSLDLMPLIPTIRETWTRKGI 265
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               Gaps
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      55 -- ESVKQLGTQNVLDLIKLVEWNH----NFGIHFMRVSSDLFPFASHAKYGYTLEFAQSHL 109
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     229 EENGSMENVCLFLNLANDPTIERIITPRIALTSAEFLAYQCKKHVLVVLTDMSSYAEALR 288
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              337 PNDDITHPIPDLTGYITEGQIYVDRQLHNRLIYP-----PINVLPSLSRLMKSAIGEGM 390
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               266 TQKQH-----YSESAD---PTAISGMKRRAHSDRVFDFPPCDPTMDLMIEAKEKE 312
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             391 TREDHSDVSNQLYACYAIGKDVQAMKAVVGEEALSSDDLLYLEFLIKFEKNFITQGHYEN 450
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          7 EMFDCLDKPI-PW------RGRLGYACLNTILRSMKERVFCSRTCRITTIQRDGL- 54
                                                                                                                                                                                                                                                                                                            SIMILARITY: BELONGS TO THE ATPASE ALPHA/BETA CHAINS FAMILY.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          66; Mismatches 158; Indels 108;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            5.1%; Score 101; DB 1; Length 491; 18.4%; Pred. No. 1.6;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              313 QAVFE-----LCRRYELQNPPCPLEIMGPEYDQTRDGYYPPGAEK 352
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       Interro; irravara...
Pfam: PF00006; ATP_synt_ab; 1.
Pfam: PF00152; ATP_synt_B_c; 1.
PROSITE; PS00152; ATPASE_ALPHA_BETA; 1.
Hydrolase; ATP synthesis; Hydrogen ion transport.
SEQUENCE 491 AA; 54750 MW; EDF4D13B3BD34716 CRC64;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          WormPep; F2086.2; CE04424.
InterPro; IPR003255; ATP_synt_B_c.
InterPro; IPR00194; ATPase_alpha_beta.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         EMBL; U41015; AAA82311.1; -.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   18.4%;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 Best Local Similarity 18.43
Matches 75; Conservative
                                           SEQUENCE FROM N.A.
                                                                    STRAIN-BRISTOL N2;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          Query Match
                                                                                           Minx P.
ρp
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      δλ
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                οy
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                q
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 δŏ
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               οy
```

Search completed: January 15, 2002, 14:07:07 Job time: 936 sec

```
Ωÿ
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             qq
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             QC
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                q
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        ^{\circ}
                                                                            Search time 57.72 Seconds (without alignments) 144.642 Million cell updates/sec
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          Appli
, Appl
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 Appl
Appli
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          Appli
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  , Appl
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     Appl
Appl
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  Appli
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                Appli
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             Appli
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            Appli
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       5183745
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         Sequence 12, Appl
Sequence 4, Appli
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   Sequence 5, Appli
Sequence 5, Appli
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              Sequence 5, Appli
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       Appli
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        Appli
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          Appli
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 Appli
                                                                                                                                                                                                                                                                                                                                                                                                                                                                 Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         Sequence 4, A
Sequence 72,
Sequence 5, A
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               Sequence 12,
Sequence 2, A
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       Sequence 2, 8 Sequence 2, 8 Sequence 14, 8 Sequence 8, 8 Sequence 8, 8 Sequence 15, 8 Sequence 15, 8 Sequence 28,
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    Sequence 2, 1
Sequence 68,
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 Sequence 4,
Sequence 4,
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                Sequence 6,
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               Description
                                                                                                                                                        .....KRLTARKRRSRKEEVEEDEK
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       Patent No.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          Sequence
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         Sequence
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               Sequence
                                                                                                                                                                                                                                                                                                                                                                          /cgn2_6/ptodata/2/iaa/5A_COMB.pep:*
/cgn2_6/ptodata/2/iaa/5B_COMB.pep:*
/cgn2_6/ptodata/2/iaa/6A_COMB.pep:*
/cgn2_6/ptodata/2/iaa/6B_COMB.pep:*
/cgn2_6/ptodata/2/iaa/PCTUS_COMB.pep:*
/cgn2_6/ptodata/2/iaa/PCTUS_COMB.pep:*
GenCore version 4.5
Copyright (c) 1993 - 2000 Compugen Ltd.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               0S-08-827-190-5
0S-08-66-858-5
0S-08-66-86-89-4
0S-08-952-127-12
0S-08-952-127-12
0S-09-103-509-4
0S-09-103-509-4
0S-09-103-66-12
0S-09-123-624-2
0S-09-125-868-15
0S-09-195-868-15
                                                                                                                                                                                                                                             hits satisfying chosen parameters:
                                                                                                                                                                                                                     212252 seqs, 22503292 residues
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        SUMMARIES
                                                                            2002, 13:51:01
                                                                                                                                                                                                                                                                                                                        Maximum Match 100%
Listing first 45 summaries
                                                    using sw model
                                                                                                                                        1979
1 DDHAPREMFDCLDKPIPWRG..
                                                                                                                                                                               BLOSUM62
Gapop 10.0 , Gapext 0.5
                                                                                                                                                                                                                                                                                                                                                                Issued_Patents_AA:*
                                                                                                                                                                                                                                                                       seq length: 0
seq length: 2000000000
                                                                                                                                                                                                                                                                                                             Post-processing: Minimum Match 0%
                                                                                                                              US-09-724-296-4
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                Query
Match Length DB
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    976
976
976
1187
11189
1229
540
                                                    protein search,
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      800
365
793
793
3066
773
773
773
773
11144
11144
                                                                            January 15,
                                                                                                                                                                                                                                             Total number of
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      888.55
88.55
88.55
88.55
88.55
88.55
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         84.5
84.5
84.5
84
                                                                                                                                          Perfect score:
                                                                                                                                                                               Scoring table:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            Score
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               888
888
887
887
887
887
887
887
                                                                                                                                                                                                                                                                     Minimum DB s
Maximum DB s
                                                    OM protein
                                                                                                                                                        Sequence:
                                                                                                                                                                                                                     Searched:
                                                                                                                                                                                                                                                                                                                                                               Database
                                                                            ..
oo
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               Result
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           Š
```

```
13, Appl
2, Appl
12, Appl
12, Appl
12, Appl
8, Appl
3, Appl
3, Appl
16, Appl
16, Appl
46, Appl
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        79 GIHFMRVSSDLFPFASHAKYGYTL--EFAQSHLEEVGKLANKYNHRLTMHPGQYTQIASP 136
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          137 REVVVDSAIRDLAYHDEILSRMKLN----EQLNKDAVLIIHLGGTFEGKKETLDRFRKNY 192
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 434 LLESNNQVYEFRISDENNEVQYKTKEGKITVLGEKFN-----WRNIEVMAKNVEGVLKP 487
  Sequence 13, R
Sequence 2, Al
Sequence 2, Al
Sequence 12, R
                                                                                     Sequence 3, Al
Sequence 3, Al
Sequence 3, Al
Sequence 3, Al
Sequence 16, A
                                                                                                                                                                          Sequence 46,
Sequence 38,
Sequence 2, A
Sequence 2, A
                                                   Sequence 12,
                                                                                                                                                  Sequence 46,
Sequence 46,
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           244 - EGSLDLMPLTPTIRETWTRKGITQKQHYSESADPTAISGMKRRAHSDRVFDFPPCDPTM
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           193 QRLSDS--VKARLVLENDDVSWSVQD--LLPLCQELNIPLVLDWHHHNI----VPGTLR-
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         Sequence Sequence 8
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               US-08-827-190-5
Sequence 5, Application US/08827190
Patent No. 5858367
GENERAL INFORMATION:
APPLICANT: Rather, Philip N.
TITLE OF INVENTION: Utilizing aarC And Compositions Thereof
NUMBER OF SEQUENCES: 13
                                                                                                                                                                                                                                                                                                                            PHILLIPPE; KRIN, EVELYN;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         5.1%; Score 101; DB 6; Length 800;
21.4%; Pred. No. 0.071;
                                                                                                                                                                                                                                                                                                                          ; APPLICANT: DANCHIN, ANTOINE;GLASER, PHILLIPPE;KRIN, EVEL);BARZU, OCTAVIEN;LADANT, DANIFL;ULLMAN, AGNES
TITLE OF INVENTION: ADENYL CYCLASE DERIVATIVES AND THEIR
US-09-052-778-13
US-08-286-325A-2
US-08-390-874C-12
US-08-390-874C-12
US-08-15-772-12
US-08-15-772-13
US-08-15-772-13
US-08-15-772-13
US-08-15-772-13
US-08-15-772-13
US-08-15-772-13
US-08-14-13-13
US-08-14-13-14-14
US-08-146-245A-46
US-08-146-245A-46
US-08-188B-46
US-08-286-188B-46
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   Mismatches
                                                                                                                                                                                                                                                   ALIGNMENTS
                                                                                                                                                                                                                                                                                                                                                                                   CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/426,541
FILING DATE: 25-OCT-1989
;SEQ ID NO:3:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          521 ----NSLEKQKGVT'NLLIKYGIERKP 542
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           DLMIEAKEKEQAVFELCRRYELQNPP 328
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                47;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     21.4%;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   Conservative
  4440
7227
7223
30566
30566
3056
3056
281
281
4823
4811
                                                                                                                                                                                                                                                                                                                                                             BIOLOGICAL USES

NUMBER OF SEQUENCES: 13
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   Best Local Similarity
Matches 57; Conserv
   Patent No. 5183745
                                                                                                                                                                                                                                                                                                                                                                                                                                         LENGTH: 800
 82
82
80
80
80
80
80
80
80
80
79
79
79
79
79
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          Query Match
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           303
                                                                                                                                                                                                                                                                                                   5183745-3
                                                                                                                                                                                                                                                                                                                                                                                                                                                   5183745-3
 QQ
```

```
TELEFAX:
                                                 COUNTRY:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             US-08-676-444-5
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             US-08-468-558-5
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  LENGTH:
                           STATE:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          TYPE:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    RESULT
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    qq
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     ŏ
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           ŏ
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       QQ
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               q
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              δλ
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            4.8%; Score 95; DB 2; Length 365;
20.4%; Pred. No. 0.086;
tive 53; Mismatches 117; Indels 114; Gaps
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      47 QIQSLERVGVDIVRVSVPTMDAAEAFKLIKQRVNVPLVADIHFDYRIAMKV-----AEY 100
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  99 GYTLEFAQSHLEEVGKLANKYNHRLTMHPGQYTQIASPREVVVDSAIRDLAYHDEILSRM 158
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             159 KLNEQLNKDAVLIIHLGGTFEGKKETLDRF---------RKNYQRLSDS 198
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     199 VKARLVLENDDVSWSVQDLLPLCQELNIPLVLDWHHHNIV-PGTLREGS----LDLMPLI 253
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               180 VKA-----SDVFLAVGSYRLLAQKIDQPL----HLGITEAGGARSGSVKSAIGLGMLL 228
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 229 -----AEGIGDTLRISLAADPVEEVKVGFDILKSLRIRSRGINFIACPTCSRQEFDV 280
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              101 GVDC-----DSARHHNIPIRI 134
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         135 GVN-------GGSLE--KDIQEKYGEPTPEALVESAMRHVDILDRLNFDQFKVS 179
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  254 PTIRETWTRKGITQKQHYSESADPT-----AISGMKRRAHSDRVFDFPPCD-PTMDL 304
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          45 RITTIQRDGLESVK-QLGTQNVLDLIKLVEWNHNF ---- GIHF-MRVSSDLFPFASHAKY 98
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           305 MIEAKEKEQAVFELCRRYELQ-----NPPCPLEI--MGPEYDQTRDGYYPPGAEKR 353
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        281 IGTVNALEQRLEDIITPMDVSIIGCVVNGPGEAEVSTLGVAGAKTKSGFYEDGVRKK 337
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                TITLE OF INVENTION: Cloning and Expression of Thermostable TITLE OF INVENTION: MutS Genes and Proteins and Uses Therefor
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              E: Hamilton, Brook, Smith & Reynolds, P.C.
Two Militia Drive
                                                                                                                                                                                E: Floppy disk
IBM PC compatible
YSTEM: PC-DOS/MS-DOS
PatentIn Release #1.0, Version #1.30
CORRESPONDENCE ADDRESS:
ADDRESSEE: Medlen & Carroll, LLP
STREET: 220 Montgomery Street, Suite 2200
                                                                                                                                                                                                                                                                                                                                                                                              NAME: Carroll, Peter G.
REGISTRATION NUMBER: 32,837
REFERENCE/DOCKET NUMBER: CASE-02443
TELECOMMUNICATION INFORMATION:
                                                                                                                United States of America
                                                                                                                                                                                                                                                                                                      APPLICATION NUMBER: US/08/827,190
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     Sequence 5, Application US/08468558
Patent No. 5877280
GENBEAL INFORMATION:
APPLICANT: Wetmur, James G.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  TELEPHONE: (415) 705-8410
TELEFAX: (415) 397-8338
INFORMATION FOR SEQ ID NO: 5:
SEQUENCE CHARACTERISTICS:
LENGTH: 365 amino acids
                                                                                                                                                                                                                                                                                                                                                                        ATTORNEY/AGENT INFORMATION:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            Query Match
Best Local Similarity 20.4%
Matches 73; Conservative
                                                                                                                                                                                                                                                                              CURRENT APPLICATION DATA:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    38
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              TOPOLOGY: linear
MOLECULE TYPE: protein
US-08-827-190-5
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     CORRESPONDENCE ADDRESS:
ADDRESSEE: Hamilton,
                                                                                                                                                               COMPUTER READABLE FORM: MEDIUM TYPE: Floppy
                                                                       San Francisco
                                                                                                                                                                                                                                                                                                                               FILING DATE:
CLASSIFICATION: 435
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    NUMBER OF SEQUENCES:
                                                                                          California
                                                                                                                                                                                                         COMPUTER: IBM PC
OPERATING SYSTEM:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          amino acid
GY: linear
                                                                                                                                         94104
                                                                                                                   COUNTRY:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               US-08-468-558-5
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    STREET:
                                                                                             STATE:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        엄
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  δλ
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              qq
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             δλ
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     δ
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 PP
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            Dp
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             δλ
```

```
292 LHPLVDR-----KQIEERLKAVERLVNDRVSLEEMRNLLSNVRDVE-RIVSRVEYNRSVP 345
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      130 YCTVFCDVSTGEVLVHESSDE---QETLDLLK-----NYSISQIICPEHLKSSLKERFF 180
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      92 FASHAKYGYTL-EFAQSHLEEVGKLAN-KYNHRLTMHP--------GQYTQIA--- 134
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               135 -----SPRE-VVVDSA-IRDLAYHDEILSRMKLNEQLNKDAVLIIHLGGTFEGKKET 184
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    237 NLKPPLLISQRDYMILDSATVENLSWIPGDRGK-NLFDVLNNTETPM----GARLGKKWI 291
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            185 L----DRFRKNYQRLSDSVKARLVLENDDVSW-SVQDLLPLCQELNIPLVLDWHHHNIVP 239
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      240 G---TLREGSLDLMPLIPTIRETWTRKGITQKQHYSESA-----DPTAISG---- 282
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              346 RULVALRE----TLEIIPKLNEVLSTFGVFKKLAFPEGLVDLLRKAIEDDPVGSPGEGKV 401
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     283 MKRRAHSDRVFDFPPCDPTMDLM-----IEAKEKEQAVFELCR-----RYELQN 326
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    64; Mismatches 138; Indels 114; Gaps
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         39 FCSRTCRITTIQRDGLESVKQLGTQNVLDLIKLVEWNHNFGI------HFMRVSSDLFP 91
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             APPLICANT: Wetmur, James G.
TITLE OF INVENTION: CLONING AND EXPRESSION OF THERMOSTABLE
TITLE OF INVENTION: MUTL GENES AND PROTEINS AND USES THEREFOR
FILE REFERENCE: MSM95-02
CURRENT APPLICATION NUMBER: US/08/676,444A
CURRENT FILING DATE: 1996-07-05
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                327 PPCPLEIMGPEYDQTR-----DGYYPPGAEKRLTARKRRSRKEEVEEDEK 371
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        ouery Match
4.5%; Score 90; DB 2; Length 793;
Best Local Similarity 22.9%; Pred. No. 1.1;
Matches 94; Conservative 64; Mismatches 138; Talli
                                                                                                                                 COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentin Release #1.0, Version #1.30
                       United States of America
                                                                                                                                                                                                                                             US/08/468,558
                                                                                                                                                                                                                                                                                                                                                                                                                     MSM94-05
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  Sequence 5, Application US/08676444A Patent No. 6294325
                                                                                                                                                                                                                                                                                                                                                                                       32,227
                                                                                                                                                                                                                                                                                                                                                                                                            REFERENCE/DOCKET NUMBER: MS
TELECOMMUNICATION INFORMATION:
                                                                                                                                                                                                                                                            06-JUN-1995
Nr: 435
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         INFORMATION FOR SEQ ID NO: 5:
SEQUENCE CHARACTERISTICS:
                                                                                                      Floppy disk
                                                                                                                                                                                                                                                                                                                                ATTORNEY/AGENT INFORMATION:
                                                                                                                                                                                                                                                                                                                                                       NAME: Granahan, Patricia
REGISTRATION NUMBER: 32.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                 617-861-6240
617-861-9540
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   : 793 amino acids
amino acid
                                                                                                                                                                                                                   CURRENT APPLICATION DATA:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   MOLECULE TYPE: protein
Massachusetts
                                                                            COMPUTER READABLE FORM: MEDIUM TYPE: Floppy
                                                                                                                                                                                                                                             APPLICATION NUMBER:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           linear
                                                                                                                                                                                                                                                                                                     CLASSIFICATION:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   GENERAL INFORMATION;
                                                                                                                                                                                                                                                                           FILING DATE:
                                              02173
                                                                                                                                                                                                                                                                                                                                                                                                                                                                       TELEPHONE:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           TOPOLOGY:
```

```
QQ
                                                                                                                                                                                                                                                                                                                                                                                                                                                                ΟŊ
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     qq
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  δλ
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   δŏ
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   ò
                                                                                                                                                                                                                             25;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  283 MKRRAHSDRVFDFPPCDPTMDLM-----IEAKEKEQAVFELCR-----RYELQN 326
                                                                                                                                                                                                                                                                                                              130 YCTVFCDVSTGEVLVHESSDE---QETLDLLK-----NYSISQIICPEHLKSSLKERFP 180
                                                                                                                                                                                                                                                                                                                                                                             181 ----GVYTETISEWYFSDLEEVERAYNLKDIHHFELSPLALKALAALIKYVKYTMIGEDL 236
                                                                                                                                                                                                                                                                                                                                                                                                                                          -----SPRE-VVVDSA-IRDLAYHDEILSRMKLNEQLNKDAVLIIHLGGTFEGKKET 184
                                                                                                                                                                                                                                                                                                                                                                                                                                                                  185 L----DRFRKNYQRLSDSVKARLVLENDDVSW-SVQDLLPLCQELNIPLVLDWHHHNIVP 239
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           240 G---TLREGSLDLMPLIPTIRETWTRKGITQKQHYSESA------DPTAISG---- 282
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    346 RDLVALRE----TLEIIPKLNEVLSTFGVFKKLAFPEGLVDLLRKAIEDDPVGSPGEGKV 401
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       402 IKRGFSSE-----LDEYRDLLEHAEERLKEFEEKERERTGIOKLRVGYNQVFGYYIEV 454
                                                                                                                                                                                                                             Gaps
                                                                                                                                                                                                                                                                       ----HFMRVSSDLFP 91
                                                                                                                                                                                  4.5%; Score 90; DB 4; Length 793; ilarity 22.9%; Pred. No. 1.1; Conservative 64; Mismatches 138; Indels 114;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                327 PPCPLEIMGPEYDQTR-----DGYYPPGAEKRLTARKRRSRKEEVEEDEK 371
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        455 TKANLDKIPDDYERKQTLVNSERFITP--ELKEFETKIMAAKERIEELEK 502
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    ADDRESSEE: Kohn & Associates
STREET: 30500 No. 6211336thwestern Hwy., Suite 410
CITY: Farmington Hills
                                                                                                                                                                                                                                                                       39 FCSRTCRITTIQRDGLESVKQLGTQNVLDLIKLVEWNHNFGI---
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               SOFTWARE: Patentin Release #1.0, Version #1.30 CURRENT APPLICATION DATA:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    APPLICANT: Shiloh, Yosef
APPLICANT: Tagle, Danilo A.
APPLICANT: Collins, Francis S.
NUMBER OF INVENTION: ATAXIA-TELANGIECTASIA GENE
NUMBER OF SEQUENCES: 24
                                                                                                                                                                                                                                                                                                                                                       92 FASHAKYGYTL-EFAQSHLEEVGKLAN-KYNHRLTMHP----
                 FastSEQ for Windows Version 4.0
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 ATTORNEY/AGENT INFORMATION:
NAME: Kohn, Kenneth I.
REGISTRATION NUMBER: 30,995
REFERENCE/DOCKET NUMBER: 2290.00029
TELECOMMUNICATION INFORMATION:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        APPLICATION NUMBER: US/08/952,12
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             PC-DOS/MS-DOS
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            Sequence 12, Application US/08952127 Patent No. 6211336 GENERAL INFORMATION:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     IBM PC compatible
                                                                                           CONGANISM: Thermotoga maritima US-08-676-444-5
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  Floppy disk
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            COMPUTER READABLE FORM: MEDIUM TYPE: Floppy of
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               CORRESPONDENCE ADDRESS:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 CLASSIFICATION: 530
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             OPERATING SYSTEM:
NUMBER OF SEQ ID NOS:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               STATE: Michigan COUNTRY: U.S.
                                                                                                                                                                                    Query Match
Best Local Similarity
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              FILING DATE:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      48334
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        COMPUTER:
                                                                                                                                                                                                                         94;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        US-08-952-127-12
                                                          793
                 SOFTWARE:
SEQ ID NO 5
                                                                              TYPE: PRT
                                                            LENGTH
                                                                                                                                                                                                                             Matches
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       qq
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     qq
                                                                                                                                                                                                                                                                                                                                                       õ
                                                                                                                                                                                                                                                                                                                                                                                              g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                               g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              ò
                                                                                                                                                                                                                                                                                                                                                                                                                                          ò
```

```
11;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       --- GKKETLDRFRKNYQ 193
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                194 RLSD-SVKARLVLENDDVSWSVQDLLPLCQELNIPLVLDWHHHNIVPGTLREGSLDLMPL 252
                                                                                                                                                                                                                                                                                                                                                                                                   87 SDLFPFASHAKYGYTLEFAQSHLEEVG---KLANKYNHRLTMHPGQYTQIASPREVVVDS 143
                                                                                                                                                                                                                                                                                        Gaps
                                                                                                                                                                                                                                                                                                                          41 SRTCRITTIQ-RDGLESVKQ--LGT-----QNVLDLIKLVEWNHNFGIHFMRVS 86
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   TREHALOSE PHOSPHORYLASE, ITS PREPARATION AND USES
                                                                                                                                                                                                                                                    4.5%; Score 89.5; DB 4; Length 3066; 22.5%; Pred. No. 12;
                                                                                                                                                                                                                                                                                        Indels
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       SOFTWARE: Patent In Release #1.0, Version #1.30 CURRENT APPLICATION DATA: APPLICATION NUMBER: US/08/966,389 FILING DATE: 07-NOY-1997
                                                                                                                                                                                                                                                                                        45; Mismatches 125;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                           144 AIRDLAYH-DEILSRMKLNEQLNKDAVLIIHLGGTFE----
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  3: BROWDY AND NEIMARK
419 Seventh Street, N.W., Suite 300
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         PRIOR APPLICATION DATA:
APPLICATION NUMBER: JP 311,232/1996
FILING DATE: 8-NOV-1996
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              1739 AATCLKNILATKIGHIFWENYKTSADP 1765
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        253 IPTIRET -- WTRKGITQKQHYSESADP 277
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            E: Floppy disk
IBM PC compatible
SYSTEM: PC-DOS/MS-DOS
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                Sequence 4, Application US/08966389
Patent No. 5843748
GENERAL INFORMATION:
               TELEFAX: 810-539-5055
INFORMATION FOR SEQ ID NO: 12:
SEQUENCE CHARACTERISTICS:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     APPLICANT: Tetsuya NAKADA
APPLICANT: Michlo KUBOTA
APPLICANT: Hiroto CHAEN
APPLICANT: Toshio MIYAKE
TITLE OF INVENTION: TREHALO
TITLE OF INVENTION: TND USE
NUMBER OF SEQUENCES: 10
                                                                   3066 amino acids
810-539-5050
                                                                                                                                                                                ORGANISM: Mus musculus
                                                                                                                                                                                                                                                                                        60; Conservative
                                                                                                           single
                                                                                                                                           MOLECULE TYPE: protein ORIGINAL SOURCE:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           COMPUTER READABLE FORM:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 CORRESPONDENCE ADDRESS:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                OPERATING SYSTEM:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      CITY: Washington
                                                                                          amino acid
                                                                                                                            linear
                                                                                                                                                                                                                                                    Query Match
Best Local Similarity
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       FILING DATE: 07 CLASSIFICATION:
                                                                                                       STRANDEDNESS:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            MEDIUM TYPE:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        20004
 TELEPHONE:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  ADDRESSEE:
                                                                                                                            TOPOLOGY:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 COMPUTER:
                                                                                                                                                                                                 US-08-952-127-12
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 US-08-966-389-4
                                                                     LENGTH:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         COUNTRY:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      STREET:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     STATE:
                                                                                                                                                                                                                                                                                        Matches
```

```
CLASSIFICATION:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        RESULT
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         q
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        οy
                                                                                                                                                                                                                                                                                                                                                                                                                                                                        461 YGCEILFETARFWEDLGAYIPLKGNKFCINCVTGPDEYTALVDNNAYTNYMAKMNLEYAY 520
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             521 DIANKM-----VASKLNLKDEEIVA 550
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     | : || :|| || : || 551 WKKAADNMYLPYSKELDIIPQDDSFLYKERITVDEIPEDQFPLLLHWHYLNIYRYQICK- 609
                                                                                                                                                                                                                                                                                                                                                                                                                                 98 YGYTLEFAQSHL-EEVGKL----ANKYNHRLIMHPGQYTQIASPREVVVDSAIRDLAYHD 152
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      153 EILSRMKLNEQLNKDAVLIIHLGGTFEGKKETLDRFRKNYQRLSDSVKARLVLENDD-VS 211
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              -----PLVLDWHHHNIVPGTLREG 245
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              246 SLDLMPLIPTIRETWTRKGITQKQHYSE-----SADPTAISGMKRR-AHSDRVFDFPP 297
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       TITLE OF INVENTION: TREHALOSE PHOSPHORYLASE, ITS PREPARATION TITLE OF INVENTION: AND USES
TUTLE OF SEQUENCES: 10
CORRESPONDENCE ADDRESS:
                                                                                                                                                                                                                                                                                                                                            Query Match 4.5%; Score 88.5; DB 2; Length 773; Best Local Similarity 20.2%; Pred. No. 1.5; Matches 50; Conservative 41; Mismatches 85; Indels 72
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patent In Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                E: BROWDY AND NEIMARK 419 Seventh Street, N.W., Suite 300
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           PRIOR APPLICATION DATA:
APPLICATION NUMBER: 08/966,389
FILING DATE: 07-NOV-1997
APPLICATION NUMBER: JP 311,232/1996
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         212 W---SVQDLLPLCQELNI-------
                                                                NAKADA=6
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             APPLICATION NUMBER: US/09/103,509
                           REGISTRATION NUMBER: 25,618
REFERENCE/DOCKET NUMBER: 26,618
TELECOMMUNICATION INFORMATION:
TELEPHONE: 202-628-5197
TELEPHONE: 202-737-3528
INFORMATION FOR SEQ ID NO: 4:
SEQUENCE CHARACTERISTICS:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    Sequence 4, Application US/09103509 Patent No. 5876975
ATTORNEY/AGENT INFORMATION:
                                                                                                                                                                                              : 773 amino acids amino acid
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  APPLICANT: Tetsuya NAKADA
APPLICANT: Michio KUBOTA
APPLICANT: Hiroto CHAEN
APPLICANT: Toshio MIYAKE
                    BROWDY, Roger L.
                                                                                                                                                                                                                                                 ; MOLECULE TYPE: peptide US-08-966-389-4
                                                                                                                                                                                                                                          linear
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            Washington
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  GENERAL INFORMATION:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           CLASSIFICATION:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               670 MTARMDL 676
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   298 CDPTMDL 304
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       USA
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                STATE: D.C
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      FILING DATE:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     COUNTRY: US
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     ADDRESSEE:
                                                                                                                                                                                                                                     TOPOLOGY:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          RESULT 7
US-09-103-509-4
                                                                                                                                                                                              LENGIH:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       STREET:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                           g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      QΥ
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             pp
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         δ
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     q
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     d
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   ŏ
```

```
10;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          Query Match 4.5%; Score 88.5; DB 2; Length 773; Best Local Similarity 20.2%; Pred. No. 1.5; Matches 50; Conservative 41; Mismatches 85; Indels 71; Gaps
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        98 YGYTLEFAQSHL-EEVCKL----ANKYNHRLTMHPGQYTQIASPREVVVDSAIRDLAYHD 152
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         461 YGCEILFETARFWEDLGAYIPLKGNKFCINCVTGPDEYTALVDNNAYTNYMAKMNLEYAY 520
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        153 EILSRMKLNEQLNKDAVLIIHLGGTFEGKKETLDRFRKNYQRLSDSVKARLVLENDD-VS 211
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         521 DIANKM-----VASKLNLKDEEIVA 550
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           212 W---SVQDLLPLCQELNI-----------PLVLDWHHHNIVPGTLREG 245
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        246 SLDLMPLIPTIRETWTRKGITQKQHYSE-----SADPTAISGMKRR-AHSDRVFDFPP 297
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          610 QPDVLLLMFLQREKFTKDELKKNYDYYEPITTHDSSLSPAIFSILANEIGYTDKAYKYFM 669
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            GENERAL INFORMATION:
APPLICANT: TetSUYa NAKADA
APPLICANT: Michio KUBOTA
APPLICANT: Hiroto CHAEN
APPLICANT: Toshio MYAKE
TITLE OF INVEXTION: TREHALOSE PHOSPHORYLASE, ITS PREPARATION
TITLE OF INVEXTION: AND USES
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patent In Release #1.0, Version #1.30
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 419 Seventh Street, N.W., Suite 300
               PRIOR APPLICATION DATA:
APPLICATION NUMBER: JP 61,716/1997
FILING DATE: 3 MAR-1997
ATTORNEY/AGENT INFORMATION:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        APPLICATION NUMBER: US/09/102,644
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             Sequence 4, Application US/09102644 Patent No. 5910436
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            BROWDY AND NEIMARK
                                                                                                                                          REGISTRATION NUMBER: 25,618
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           IBM PC compatible
                                                                                                                                                                                              TELECOMMUNICATION INFORMATION:
                                                                                                                                                                                                                                             TELEFAX: 202-737-3528
INFORMATION FOR SEQ ID NO: 4:
SEQUENCE CHARACTERISTICS:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      Floppy disk
                                                                                                                                                                    REFERENCE/DOCKET NUMBER:
                                                                                                                                                                                                                   : 202-628-5197
202-737-3528
FILING DATE: 8-NOV-1996
                                                                                                                                                                                                                                                                                                                      773 amino acids
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              CURRENT APPLICATION DATA:
                                                                                                                          Roger L.
                                                                                                                                                                                                                                                                                                                                                                                       ; MOLECULE TYPE: peptide US-09-103-509-4
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            COMPUTER READABLE FORM:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        NUMBER OF SEQUENCES: 1
CORRESPONDENCE ADDRESS:
                                                                                                                                                                                                                                                                                                                                             TYPE: amino acid
                                                                                                                                                                                                                                                                                                                                                                       linear
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           Washington
                                                                                                                     BROWDY,
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     670 MTARMDL 676
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       298 CDPTMDL 304
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      MEDIUM TYPE:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  FILING DATE:
                                                                                                                                                                                                                   TELEPHONE:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  20004
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            ADDRESSEE:
                                                                                                                                                                                                                                                                                                                                                             TOPOLOGY:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           COMPUTER:
                                                                                                                                                                                                                                                                                                                      LENGTH:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      US-09-102-644-4
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             COUNTRY:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      STREET:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   STATE:
```

```
q
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    QQ
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 qq
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       ŏ
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                Qy
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                10;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             98 YGYTLEFAQSHL-EEVGKL----ANKYNHRLTMHPGQYTQIASPREVVVDSAIRDLAYHD 152
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       461 YGCEILFETARFWEDLGAYIPLKGNKFCINCVTGPDEYTALVDNNAYTNYMAKMNLEYAY 520
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    EILSRMKLNEQLNKDAVLIIHLGGTFEGKKETLDRFRKNYQRLSDSVKARLVLENDD-VS 211
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              521 DIANKM-----VASKLNLKDEEIVA 550
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        212 W---SVQDLLPLCQELNI----------PLVLDWHHHNIVPGTLREG 245
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    551 WKKAADNMYLPYSKELDIIPQDDSFLYKERITVDEIPEDQFPLLLHWHYLNIYRYQICK- 609
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 246 SLDLMPLIPTIRETWTRKGITQKQHYSE-----SADPTAISGMKRR-AHSDRVFDFPP 297
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           610 QPDVLLLMFLQREKFTKDELKKNYDYYEPITTHDSSLSPAIFSILANEIGYTDKAYKYFM 669
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     GENERAL INFORMATION:
APPLICANT: Tetsuya NAKADA
APPLICANT: Michio KUBOTA
APPLICANT: HICCOC CHAEN
APPLICANT: Toshio MIYAKE
TITLE OF INVENTION: TREBALOSE PHOSPHORYLASE, ITS PREPARATION
TITLE OF INVENTION: TABU USES
NUMBER OF SEQUENCES: 10
CORRESPONDENCE ADDRESS:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    Length 773;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               COMPUTER: IBM PC Compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patent In Release #1.0, Version #1.30
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 Query Match
Best Local Similarity 20.2%; Pred. No. 1.5;
Matches 50; Conservative 41; Mismatches 85;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               ADDRESSEE: BROWDY AND NEIMARK
STREET: 419 Seventh Street, N.W., Suite 300
CITY: Washington
                                                               FILING DATE: 07-NOV-1997
APPLICATION NUBBER: JP 311,232/1996
FILING DATE: 8-NOV-1996
PRIOR APPLICATION DATA:
APPLICATION NUBBER: JP 61,716/1997
FILING DATE: 3-MAR-1997
ATTORNEY/AGENT INFORMATION:
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 08/966,389
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           Sequence 4, Application US/09218032
Patent No. 5993889
                                                                                                                                                                                                                REGISTRATION NUMBER: 25,618
REGISTRATION NUMBER: 25,618
REFERENCE/DOCKET NUMBER: NAM
TELECOMMUNICATION
TELEPHONE: 202-628-5197
                                                                                                                                                                                                                                                                                                                                                          INFORMATION FOR SEQ ID NO: 4: SEQUENCE CHARACTERISTICS:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
                                                                                                                                                                                                                                                                                                                                                                                                         : 773 amino acids amino acid
                                                                                                                                                                                                                                                                                                                                   202-737-3528
                                                                                                                                                                                                                                                                                                                                                                                                                                                                             MOLECULE TYPE: peptide
                                                                                                                                                                                                                                                                                                                                                                                                                                                          linear
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  670 MTARMDL 676
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        298 CDPTMDL 304
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   STATE: D.C.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               20004
                                                                                                                                                                                                                                                                                                                                                                                                                                                   TOPOLOGY:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        COUNTRY:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         US-09-218-032-4
                                                                                                                                                                                                                                                                                                                                                                                                         LENGTH:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  US-09-102-644-4
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    Ω
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    ò
```

```
98 YGYTLEFAQSHL-EEVGKL----ANKYNHRLTMHPGQYTQIASPREVVVDSAIRDLAYHD 152
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       461 YGCEILFETARFWEDLGAYIPLKGNKFCINCVTGPDEYTALVDNNAYTNYMAKMNLEYAY 520
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               153 EILSRMKLNEQLNKDAVLIIHLGGTFEGKKETLDRFRKNYQRLSDSVKARLVLENDD-VS 211
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 521 DIANKM------VASKLNLKDEEIVA 550
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       ---PLVLDWHHHNIVPGTLREG 245
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      | : || :|| || : |
551 WKKAADNMYLPYSKELDIIPQDDSFLYKERITVDEIPEDQFPLLLHWHYLNIYRYQICK• 609
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               246 SLDLMPLIPTIRETWTRKGITQKQHYSE-----SADPTAISGMKRR-AHSDRVFDFPP 297
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  610 QPDVLLLMFLQREKFTKDELKKNYDYYEPITTHDSSLSPAIFSILANEIGYTDKAYKYFM 669
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 71; Gaps
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                TITLE OF INVENTION: No. 6248562el Chimeric Proteins Comprising TITLE OF INVENTION: Borrelia Polypeptides and Uses Therefor
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      Query Match 4.5%; Score 88.5; DB 2; Length 773; Best Local Similarity 20.2%; Pred. No. 1.5; Matches 50; Conservative 41; Mismatches 85; Indels 7.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATION SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentin Release #1.0, Version #1.25
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              Brookhaven National Laboratory
                                                                                                                                                                           PRICE APPLICATION DATA:
APPLICATION NUMBER: JP 61,716/1997
FILING DATE: 3-MAR-1997
ATTORNEY/AGENT INFORMATION:
                                                                                                                                                                                                                                                                                                                                               REFERENCE/DOCKET NUMBER: NAKADA=6
                      APPLICATION NUMBER: US/09/218,032
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   Sequence 72, Application US/08235836C Patent No. 6248562 GENERAL INFORMATION:
                                                                                                        PRIOR APPLICATION DATA:
APPLICATION NUMBER: 08/966,389
FILING DATE:
                                                                                                                                                                                                                                                                                                                        25,618
                                                                                                                                                                                                                                                                                                                                                                           TELECOMMUNICATION INFORMATION:
TELEPHONE: 202-628-5197
TELEFAX: 202-737-3528
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       212 W---SVQDLLPLCQELNI-----
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             APPLICANT: Dunn, John J. APPLICANT: Luft, Benjamin J.
                                                                                                                                                                                                                                                                                                                                                                                                                                      TELEFAX: 202-737-3528
INFORMATION FOR SEQ ID NO: 4:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               LENGTH: 773 amino acids
                                                                                                                                                                                                                                                                                            NAME: BROWDY, ROGER L. REGISTRATION NUMBER: 2:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         SEQUENCE CHARACTERISTICS:
CURRENT APPLICATION DATA:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               TOPOLOGY: linear

MOLECULE TYPE: peptide

US-09-218-032-4
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         COMPUTER READABLE FORM: MEDIUM TYPE: Floppy
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      NUMBER OF SEQUENCES: 1 CORRESPONDENCE ADDRESS:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 TYPE: amino acid
                                                                              CLASSIFICATION:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       670 MTARMDL 676
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       298 CDPTMDL, 304
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         USA
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   Upton
                                                     FILING DATE:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      COUNTRY: U
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              ADDRESSEE:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     US-08-235-836C-72
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             Ν
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          STATE:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               RESULT 10
```

us-09-724-296-4.rai

9

```
Qγ
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   qq
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         qq
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         δò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  18;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     105 AQSHLEEVGKLANKYNHRLTMHPGQYT-QIASPREVVVDSAIRDLAYHDEILSRMKLNEQ 163
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             179 AGKTQIFIPLKKNILSGKIESDIDIDSL-----VTDKVVAALLSENEAGVNFARDIT-D 231
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   214 VQDLLPLCQELNIPLVLDWHHHNIVPGTLREGSLDLMPLIPTIRETWTRKGITQKQHYSE 273
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        71 LVEWNHN---FGIHFM------RVSSDLFPFASHAKY-------GY---TLEF 104
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           164 LNKDAVLIIHLGGTFEGKKET---LDRFRKNYQRLSDSVKARLVLEND-----DVSWS 213
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            232 IQGETHKADQDKIDIELD-----NVHESDSNITETIENLRDQLEK--ATDEEHRKE 280
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       274 SADPTAISGMKRRAHSDRVFDFPPCDPTMDLMIE-----AKEK-EQAVFELCRRYEL 324
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  80; Gaps
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  281 IESQVDAKKKOKEELDKKAIDLDKAQQKLDFSEDNLDIQRDTVREKIQEDINEINKE--- 337
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      APPLICANT: Xie, Qiao-wen
APPLICANT: Nathan, Carl F.
APPLICANT: Mathan, Carl F.
APPLICANT: Mumford, Richard A.
APPLICANT: Calaycay, Jimmy Ramos
APPLICANT: Calaycay, Jimmy Ramos
NUMBER OF SEQUENCES: 6
CORRESPONDENCE ADDRESS:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  Ouery Match 4.4%; Score 88; DB 4; Length 693; Best Local Similarity 17.9%; Pred. No. 1.4; Matches 62; Conservative 72; Mismatches 133; Indels
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               325 QNPPCPLEIMGPEYD---QTRDGYYPPGAEKRLTARKRRSRKEEVEE 368
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/148,191
FILING DATE: 01-11-93
ATTORNEY/AGENT INPORMATION:
NAME: BOGOSIAN, MARGARET C.
REGISTATION NUMBER: 25,324
REFERENCE/DOCKET NUMBER: BNL93-28A
TELECOMMUNICATION INPORMATION:
                APPLICATION NUMBER: US/08/235,836C
FILING DATE: 29-APR-1994
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              ADDRESSEE: Merck & Co., Inc.
STREET: 126 East Lincoln Avenue
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy Disk
COMPUTER: Macintosh Centris650
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     ; Sequence 5, Application US/08147812
; Patent No. 5766909
                                                                                                                                                                                                                                                        IELEPHONE: (516) 282-7338
TELEPAX: (516) 282-3729
INFORMATION FOR SEQ ID NO: 72:
SEQUENCE CHARACTERISTICS:
LENGTH: 693 amino acids
CURRENT APPLICATION DATA:
                                                                                                                                                                                                                                                                                                                                                                                                                   MOLECULE TYPE: protein
                                                                                                                                                                                                                                                                                                                                                                            amino acid
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      New Jersey
                                                                                                                                                                                                                                                                                                                                                                                                           linear
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                GENERAL INFORMATION:
                                                              CLASSIFICATION:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       USA
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              Rahway
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              07065
                                                                                                                                                                                                                                                                                                                                                                                                                                                    US-08-235-836C-72
                                                                                                                                                                                                                                                                                                                                                                                                       TOPOLOGY:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       COUNTRY:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  US-08-147-812-5
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               STATE:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              CITY:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   δŏ
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   qq
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    Óλ
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             ŏ
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      ò
```

```
80 IHFMRVSSDLFPFASHAKYGYTLEFAQSHL------EEVG------113
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              673 AACETFDVRSKHHIQIPKRFTSNATWEPQQYRLIQSPEPLDLNRALSSIHAKNVFTMRLK 732
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       733 SOONLOSEKSSRTTLLVOL--TFEGSRGPSYLPGEHLGIFPGNQTALVQGILERVVDCPT 790
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      615 YAVFGLGSSMYP--QFCAFAHDIDQKLSHLGASQLAPTGEGDELSGQEDAFRSWAVQTFR 672
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 114 -------KLANKYNHRLTMHPGQYTQIASPREVVVDSAIRDLAYHDEILSRMK 159
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    160 LNEQLNKD----AVLIIHLGGTFEGKK-----ETLDRFRKNYQRLSDSVKARL---- 203
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                562 KVVCMDQYKASTLEEEQLLLVVTSTFGNGDCPSNGOTLKK-----SLFMLRELNHTFR 614
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   -----ESVKQLGTQNVLDLIKLVEWNHNFG 79
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                92;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     Query Match 4.4%; Score 88; DB 1; Length 1144; Best Local Similarity 17.3%; Pred. No. 3.3; Matches 46; Conservative 38; Mismatches 90; Indels 9
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      APPLICANT: Tilly, Timothy P.
APPLICANT: Yil, Jerry C.
APPLICANT: Regulski, Michael
TITLE OF INVENTION: CLONING AND CHARACTERIZATION OF GENES
TITLE OF INVENTION: ASSOCIATED WITH LONG-TERM MEMORY
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    ADDRESSEE: Hamilton, Brook, Smith & Reynolds, P.C. STREET: Two Militia Drive
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  COMPUTER: IBM PC COMPALIBLE
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PALENTIN Release #1.0, Version #1.30
                                                                         APPLICATION NUMBER: US/08/147,812
FILING DATE: No. 5766909 Available
OPERATING SYSTEM: Macintosh 7.0.1
                                                                                                                                                                                                                                                                                                                            18658IA
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                204 -----VLENDDVSWSVODLLPLC 221
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                791 PHQTVCLEVLDESGSYWVKDKRLPPC 816
                                                                                                                                                                                       07/841,641
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  Sequence 12, Application US/08319866 Patent No. 5929223
                            SOFTWARE: Microsoft Word 5.1a CURRENT APPLICATION DATA:
                                                                                                                                                                                                                                                                                          REGISTRATION NUMBER: 35,403
REFERENCE/DOCKET NUMBER: 186
TELECOMMUNICATION INFORMATION:
                                                                                                                                                                                                                                                                                                                                                                               (908) 594-3905
(908) 594-4720
                                                                                                                                                                                                                                                                                                                                                                                                                                                      INFORMATION FOR SEQ ID NO: 5:
SEQUENCE CHARACTERISTICS:
LENGTH: 1144 amino acids
                                                                                                                                                                                                           FILLING DATE: 02-FEB-1992
ATTORNEY/AGENT INFORMATION:
                                                                                                                                                                                                                                                                    NAME: Wallen, John W LII
REGISTRATION NUMBER: 35,
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                37 RVFCSRTCRITTORDGL --
                                                                                                                                                           PRIOR APPLICATION DATA:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             CORRESPONDENCE ADDRESS:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      Massachusetts
                                                                                                                                  435
                                                                                                                                                                                       APPLICATION NUMBER:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 NUMBER OF SEQUENCES:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           amino acid
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            TOPOLOGY: linear
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            Patent No. 5929233
GENERAL INFORMATION:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         Lexington
                                                                                                                               CLASSIFICATION:
                                                                                                                                                                                                                                                                                                                                                                                                                             138825
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  USA
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      02173
                                                                                                                                                                                                                                                                                                                                                                               TELEPHONE:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         US-08-319-866-12
                                                                                                                                                                                                                                                                                                                                                                                                         TELEFAX:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           US-08-147-812-5
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         COUNTRY:
                                                                                                                                                                                                                                                                                                                                                                                                                                     TELEX:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   STATE:
```

```
9.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               673 AACETFDVRSKHHIQIPKRFTSNATWEDQQYRLIQSPEPLDLNRALSSIHAKNVFTMRLK 732
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          160 LNEQLNKD----AVLIIHLGGTFEGKK-----ETLDRFRKNYQRLSDSVKARL---- 203
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       562 KVVCMDQYKASTLEEEQLLLVVTSTFGNGDCPSNGQTLKK-----SLFMLRELNHTFR 614
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      80 IHFMRVSSDLFPFASHAKYGYTLEFAQSHL-------EEVG--------- 113
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              615 YAVFGLGSSMYP--QFCAFAHDIDQKLSHLGASQLAPTGEGDELSGQEDAFRSWAVQTFR 672
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    --KLANKYNHRLTMHPGQYTQIASPREVVVDSAIRDLAYHDEILSRMK 159
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   TITLE OF INVENTION: TREATMENT OF VASCULAR DISORDERS
TITLE OF INVENTION: TREATMENT OF VASCULAR DISORDERS
FILE REFERENCE: 511169-2003
CURRENT APPLICATION NUMBER: US/09/123,708
CURRENT FILING DATE: 1998-07-28
EARLIER APPLICATION NUMBER: 08/553,503
EARLIER PLIANG DATE: 1996-03-01
EARLIER PLIANG DATE: 1996-03-01
EARLIER PLIANG DATE: 1994-03-31
NUMBER OF SEQ ID NOS: 6
SOFTWARE: PATENTION VOS: 6
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            37 RVFCSRTCRITTIQRDGL-------ESVKQLGTQNVLDLIKLVEWNHNFG 79
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    92;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            Ouery Match

4.4%; Score 88; DB 2; Length 1144;
Best Local Similarity 17.3%; Pred. No. 3.3;
Matches 46; Conservative 38; Mismatches 90; Indels
                                                                                                                                                                                                                                      CSHL94-03
                      APPLICATION NUMBER: US/08/319,866
FILING DATE: 7-OCT-1994
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   791 PHQTVCLEVLDESGSYWVKDKRLPPC 816
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       204 ------VLENDDVSWSVQDLLPLC 221
                                                                                                                                                                                NAME: Granahan, Patricia
REGIZTATION NUMBER: 32,227
REFRENCE/DOCKET NUMBER: CSHLS
TELECOMMUNICATION INFORMATION:
TELEFAN: (617) 861-6240
INFORMATION FOR SEO ID NO: 12:
SEQUENCE CHARACTERISTICS:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       ; Sequence 2, Application US/09123708
; Patent No. 6146887
                                                                                                                                                             ATTORNEY/AGENT INFORMATION:
NAME: Granahan, Patricia
                                                                                                                                                                                                                                                                                                                                                                             1144 amino acids
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                GENERAL INFORMATION:
APPLICANT: SCHRADER, Juergen
APPLICANT: GOEDECKE, Axel
CURRENT APPLICATION DATA:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   ; TYPE: PRT
; ORGANISM: Cytomegalovirus
US-09-123-708-2
                                                                                     PRIOR APPLICATION DATA:
APPLICATION NUMBER:
FILING DATE:
                                                                                                                                                                                                                                                                                                                                                                                                                                             TOPOLOGY: linear
MOLECULE TYPE: protein
US-08-319-866-12
                                                                                                                                                                                                                                                                                                                                                                                                    TYPE: amino acid
STRANDEDNESS:
                                                                    CLASSIFICATION:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     SEQ ID NO 2
LENGTH: 1144
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        RESULT 13
US-09-123-708-2
                                                                                                                                                                                                                                                                                                                                                                             LENGTH:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      ŏ
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       à
```

```
9
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           APPLICANT: SCHRADER, Jurgen
APPLICANT: SCHRADER, Jurgen
APPLICANT: SCHRADER, Axel
TITLE OF INVENTION: DNA EXPRESSION VECTORS FOR USE IN THE GENE THERAPEUTIC
TITLE OF INVENTION: TREATMENT OF VASCULAR DISORDERS
FILE REFERENCE: 511169-2004
CURRENT APPLICATION NUMBER: US/09/123,624
CURRENT FILING DATE: 1996-07-28
PRIOR FILING DATE: 1996-03-01
PRIOR FILING DATE: 1996-03-01
PRIOR APPLICATION NUMBER: 4411402.8
PRIOR APPLICATION NUMBER: 4411402.8
NUMBER OF SEQ ID NOS: 6
SOFTWARE: PATENTING VOS: 6
SOFTWARE: PATENTING VOS: 6
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         6
                                                                                                                                                                                     80 1HFMRVSSDLFPFASHAKYGYTLEFAQSHL-------EEVG----------------- 113
                                                                                                                                                                                                                                615 YAVFGLGSSMYP--QFCAFAHDIDQKLSHLGASQLAPTGEGDELSGQEDAFRSWAVQTFR 672
                                                                                                                                                                                                                                                                                 114 ------KLANKYNHRLTMHPGQYTQIASPREVVVDSAIRDLAYHDEILSRMK 159
                                                                                                                                                                                                                                                                                                              673 AACETFDVRSKHHIQIPKRFTSNATWEPQQYRLIQSPEPLDLNRALSSIHAKNVFTMRLK 732
                                                                                                                                                                                                                                                                                                                                                                                                     562 KVVCMDQYKASTLEEEQLLLVVTSTFGNGDCPSNGQTLKK-----SLFMLRELNHTFR 614
                                                                                                                                                                                                                                                                                                                                                                        160 LNEQLNKD----AVLIIHLGGTFEGKK-----ETLDRFRKNYQRLSDSVKARL---- 203
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       673 AACETFDVRSKHHIQIPKRFTSNATWEPQQYRLIQSPEPLDLNRALSSIHAKNVFTMRLK 732
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       615 YAVFGLGSSMYP--QFCAFAHDIDQKLSHLGASQLAPTGEGDELSGQEDAFRSWAVQTFR 672
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   114 -------KLANKYNHRLTMHPGQYTQIASPREVVVDSAIRDLAYHDEILSRMK 159
                                                  90; Indels 92; Gaps
                                                                                              ---ESVKQLGTQNVLDLIKLVEWNHNFG 79
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            562 KVVCMDQYKASTLEEEQLLLVVTSTFGNGDCPSNGQTLKK-----SLFMLRELNHTFR 614
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             160 LNEQLNKD----AVLIIHLGGTFEGKK-----ETLDRFRKNYQRLSDSVKARL---- 203
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              37 RVFCSRTCRITTIQRDGL-------ESVKQLGTQNVLDLIKLVEWNHNFG 79
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       4.4%; Score 88; DB 4; Length 1144; 17.3%; Pred. No. 3.3;
    DB 4; Length 1144;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          80 IHFMRVSSDLFPFASHAKYGYTLEFAQSHL-------EEVG---
                                                  38; Mismatches
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  38; Mismatches
    4.4%; Score 88;
17.3%; Pred. No.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               791 PHQTVCLEVLDESGSYWVKDKRLPPC 816
                                                                                                                                                                                                                                                                                                                                                                                                                                                                 204 -----VLENDDVSWSVQDLLPLC 221
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         204 -----VLENDDVSWSVQDLLPLC 221
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 RESULT 14
US-00-123 424-2
Sequence 2, Application US/09123624
; Patent No. 6149936
                                                                                              37 RVFCSRTCRITTIORDGL----
Query Match
Best Local Similarity 17.3%
Matches 46; Conservative
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     Conservative
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           ORGANISM: Mus musculus
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       Ouery Match
Best Local Similarity
"Ans 46; Conserva
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               ; GENERAL INFORMATION:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               1144
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 US-09-123-624-2
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       SEQ ID NO 2
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     TYPE: PRT
                                                                                              Οy
                                                                                                                                       qq
                                                                                                                                                                                     δ
                                                                                                                                                                                                                              qq
                                                                                                                                                                                                                                                                              δ
                                                                                                                                                                                                                                                                                                                          QQ
                                                                                                                                                                                                                                                                                                                                                                        ŏλ
                                                                                                                                                                                                                                                                                                                                                                                                                      q
                                                                                                                                                                                                                                                                                                                                                                                                                                                                    ÓΫ
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           qq
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            Dp
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   pp
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          δλ
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   ōγ
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            QQ
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             δy
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           qq
```

```
APPLICANT: Pot, David A.
APPLICANT: Pot, David A.
APPLICANT: Williams, Lewis T.
APPLICANT: Jefferson, Anne Bennett
APPLICANT: Jefferson, Anne Bennett
APPLICANT: Majerus, Philip W.
TITLE OF INVENTION: No. 6001354el Grb2 Associating Protein and Nucleic
TITLE OF INVENTION: Acids Encoding Therefor
NUMBER OF SEQUENCES: 10
CORRESPONDENCE ADDRESS:
ADDRESSEE: Townsend and Townsend and Crew
STREET: One Market Plaza, Steuart Tower, Suite 2000
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                4.4%; Score 87; DB 3; Length 976;
20.4%; Pred. No. 3.3;
Live 53; Mismatches 130; Indels 126; Gaps
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    |:| : | : | |: | |: | |- |: |- |- |: |- |: |- |: |- |: |- |: |- |: |- |: |- |: |- |: |- |: |- |: |- |: |- |: |- |: |- |: |- |: |- |: |- |: |- |: |- |: |- |: |- |: |- |: |- |: |- |: |- |: |- |: |- |: |- |: |- |: |- |: |- |: |- |: |- |: |- |: |- |: |- |: |- |: |- |: |- |: |- |: |- |: |- |: |- |: |- |: |- |: |- |: |- |: |- |: |- |: |- |: |- |: |- |: |- |: |- |: |- |: |- |: |- |: |- |: |- |: |- |: |- |: |- |: |- |: |- |: |- |: |- |: |- |: |- |: |- |: |- |: |- |: |- |: |- |: |- |: |- |: |- |: |- |: |- |: |- |: |- |: |- |: |- |: |- |: |- |: |- |: |- |: |- |: |- |: |- |: |- |: |- |: |- |: |- |: |- |: |- |: |- |: |- |: |- |: |- |: |- |: |- |: |- |: |- |: |- |: |- |: |- |: |- |: |- |: |- |: |- |: |- |: |- |: |- |: |- |: |- |: |- |: |- |: |- |: |- |: |- |: |- |: |- |: |- |: |- |: |- |: |- |: |- |: |- |: |- |: |- |: |- |: |- |: |- |: |- |: |- |: |- |: |- |: |- |: |- |: |- |: |- |: |- |: |- |: |- |: |- |: |- |: |- |: |- |: |- |: |- |: |- |: |- |: |- |: |- |: |- |: |- |: |- |: |- |: |- |: |- |: |- |: |- |: |- |: |- |: |- |: |- |: |- |: |- |: |- |: |- |: |- |: |- |: |- |: |- |: |- |: |- |: |- |: |- |: |- |: |- |: |- |: |- |: |- |: |- |: |- |: |- |: |- |: |- |: |- |: |- |: |- |: |- |: |- |: |- |: |- |: |- |: |- |: |- |: |- |: |- |: |- |: |- |: |- |: |- |: |- |: |- |: |- |: |- |: |- |: |- |: |- |: |- |: |- |: |- |: |- |: |- |: |- |: |- |: |- |: |- |: |- |: |- |: |- |: |- |: |- |: |- |: |- |: |- |: |- |: |- |: |- |: |- |: |- |: |- |: |- |: |- |: |- |: |- |: |- |: |- |: |- |: |- |: |- |: |- |: |- |: |- |: |- |: |- |: |- |: |- |: |- |: |- |: |- |: |- |: |- |: |- |: |- |: |- |: |- |: |- |: |- |: |- |: |- |: |- |: |- |: |- |: |- |: |- |: |- |: |- |: |- |: |- |: |- |: |- |: |- |: |- |: |- |: |- |: |- |: |- |: |- |: |- |: |- |: |- |: |- |: |- |: |- |: |- |: |- |: |- |: |- |: |- |: |- |: |- |: |- |: |- |: |- |: |- |: |- |: |- |: |- |: |- |: |- |: |- |: |- |: |- |: |- |: |- |: |- |: |- |: |- |: |- |: |- |: |- |: |- |: |- |: |- |: |- |: |- |: |- |: |- |: |- |: |- |: |- |: |- |: |- |: |- |: |- |: |- |: |- |: |- |: |- |: |- |: |- |: |- |:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   129 QYTQIASPREVVVD-----SAIRDLAYHDEILS-RMKLNE-QLNKDAVLIIH--LGG 176
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           177 TFEGKKETLDRFRKNYQRLSDSVKARLVLENDDVS--------WSVQDLLPL 220
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     633 HFQGEIKLQTSQGKTREKLYDFVKT----ERDESSGPKTLKSLTSHDPMKQWEVTSRAPP 688
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               221 CQELNIPLVLDWHHHNIVPGTLREGSLDLMPLIPTIRETWTRKGITQKQHYSESADPTAI 280
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          689 CSGSSITEIINPNYMGVGP-----FGPPMP-----LHVKQTLSPDQQPTAW 729
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         281 SGMKRRAHSDRVFDFPPCDPTMDLMIEAKEKEQAVFELCRRYELQNPPCP--LEIMGPEY 338
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             82 FMRVSSDLFPFASHAKYGYTLEFAQSHLEEV------GKLANKYNHRLTMHPG 128
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         24 YACLNTILRSMKERVFCSRTCRITT - IQRDGLESVKQLGTQNVLDLIKLVEWNHNFGIH 81
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                SOFTWARE: Patentin Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/560,005
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   NAME: Dow, Karen B.
REGISTRATION NUMBER: 29,684
791 PHQTVCLEVLDESGSYWVKDKRLPPC 816
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            E: Floppy disk
IBM PC compatible
SYSTEM: PC-DOS/MS-DOS
                                                                                                                                                                                                                                      Sequence 2, Application US/08560005
Patent No. 6001354
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      REFERENCE/DOCKET NUMBER: 23:
TELECOMMUNICATION INFORMATION:
TELEPHONE: 415-326-2400
TELEBRAX: 415-326-2422
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          INFORMATION FOR SEQ ID NO: 2: SEQUENCE CHARACTERISTICS: LENGTH: 976 amino acids TYPE: amino acid
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               ATTORNEY/AGENT INFORMATION:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                --
=
--
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               Query Match 4.4%
Best Local Similarity 20.4%
Matches 79; Conservative
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            ; MOLECULE TYPE: protein US-08-560-005-2
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    COMPUTER READABLE FORM: MEDIUM TYPE: Floppy
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                San Francisco
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  CLASSIFICATION: 435
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       California
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              COMPUTER: IBM PC
OPERATING SYSTEM:
                                                                                                                                                                                                                                                                                                                                          GENERAL INFORMATION:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               USA
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          FILING DATE:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       94105
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              TOPOLOGY:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               COUNTRY:
                                                                                                                                            RESULT 15
US-08-560-005-2
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       STATE:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       qq
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     Dp
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     ΩD
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  Ор
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           δλ
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  δλ
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              qq
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             ò
```

```
Db 730 S-------YDQPPKDSPLG-------PCRG---ESPPTPPGQPPISPK-760
Qy 339 DQTRDGYYPPGAEKRLTARKRRSRKEEV 366
bb 761 ----KFLPSTANRGLPPRTOESRPSDL 783
Search completed: January 15, 2002, 13:51:04
Job time: 333 sec
```

	·		
			4

```
GenCore version 4.5
Copyright (c) 1993 - 2000 Compugen Ltd.
```

OM protein - protein search, using sw model

Run on: January 15, 2002, 14:07:26 ; Search time 45.07 Seconds
(without alignments)
239.171 Million cell updates/sec

Title: US-09-724-296-39
Perfect score: 1527
Sequence: 1 QLGLVCLTVGPEVRRTVTL.....VEAKGKEEAIAALRLMAPFK 294

Sequence: 1 OLGLVCLTVGPEVRFRTVTL....
Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 100059 seqs, 36664827 residues

Total number of hits satisfying chosen parameters: 100059

Minimum DB seq length: 0 Maximum DB seq length: 2000000000 Post-processing: Minimum Match 0% Maximum Match 100% Listing first 45 summaries

Database : SwissProt_39:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

	Description	P45864 bacillus su	_	800	P08633 rhizobium l		Q9wu42 mus musculu	Q60716 mus musculu			-		-				Q03132 saccharopol		P23852 escherichia	P05894 simian immu	٠.		_			Q10670 mycobacteri		002482 caenorhabdi	Q00037 escherichia	m	_	22	334	P13397 flavobacter
SUMMARIES	ID		LEF4_NPVOP	DHAL_PSESP	NODM_RHILV	PUAC_STRLP	NCR2_MOUSE	P4H2_MOUSE	PKNK_MYCTU	APP1_MOUSE	HEY1_CANFA	RFX5_HUMAN	RPSD_CLOAB	TRPA_MYCIT	APP1_HUMAN	DP3A_ECOLI	ERY2_SACER	COBB_RHOCA	HEPA_ECOLI	GAG_SIVM1	DNAK_STRPN	ATXB_LEIDO	SYK2_MYCTU	PCCB_RHOER	OBP_HSVBC	BLAC_MYCTU	TRPB_BACST	UN37_CAEEL	TNPA_ECOLI	KPBA_CAEEL	ISPF_MYCTU	TRPB_CAUCR	VNFK_AZOCH	NYLA_FLASK
	ength DB		457 1			199 1					304 1	616 1	378 1	271 1	650 1	1160 1	3567 1	435 1	967 1	506 1									1002 1			406 1		492 1
đ	Query Match Length	19.8	5.9	5.9	5.9	5.8	5.8	5.8	5.6	9.6	5.5	5.5	5.5	5.4	5.4	5.4	5.4	5.4	5.3	5.3	5.3	5.3	5.3	5.3	5.3	5.2	5.2	5.2	5.2	5.5	5.2	5.3	5.2	5.2
	Score	303	90.5	06	06	68	68	88	98	85.5	84.5	84.5	84	82.5	82.5	82.5	82.5	82	81.5	81	81	81	81	80.5	80.5	80	80	80	80	80	79.5	79.5	79.5	79.5
	Result No.		7	٣	4	S	9	7	σο		10	11	12	13	14	15	16	17	18	19	20	21	22	23	24	25	56	27	28	59	30	31	32	33

033277 moorella th	P34355 caenorhabdi	P97305 mus musculu	P98095 homo sapien	Q13085 homo sapien	P46829 mycobacteri	P12851 actinoplane	Q06805 bos taurus	Q03131 saccharopol	P52027 deinococcus	Q07017 streptomyce	Q9y5j3 homo sapien
A_MOOTH	P_CAEEL	3_MOUSE	2_HUMAN	1_HUMAN	B_MYCBO	A_ACTMI	1_BOVIN	1_SACER	1_DEIRA	OL56_STRAT	IEY1_HUMAN
1 SEL	1 CAC	1 NFC	1 FBL	1 COA	1 DAF	1 XYL	1 TIE	1 ERY	1 DPC	1 OL5	1 HEY
470	629	1075	1184	2346	271	393	1136	3491	926	3519	304
5.2	5.2	5.2	5.2	5.2	5.1	5.1	5.1	5.1	5.1	5.1	5.1
79	7.9	79	79	79	78.5	78.5	78.5	78.5	7.8	7.8	77.5
34	35	36	37	38	39	40	41	42	43	44	45

ALIGNMENTS

### PRT; 320 AA. ### ACSU STANDARD; PRT; 320 AA. ### 1995 Rel. 32, Created) ### 1995 Rel. 32, Created) ### 1995 Rel. 32, Created) ### 1995 Rel. 32, Last sequence update) ### 1996 Rel. 36, Last annotation update ### 1997 Rel. 36, Rat annotation update ### 1998 Rel. 36, Rel.	MAYD. BACSU TO AND PAGE TO AND PAGE TO AND
SU STANDA; 995 (Rel. 32, 995 (Rel. 32, 101 (Rel. 40, 102 (Rel. 32, 103 (Rel. 32, 103 (Rel. 36, 104 (Rel. 36, 105 (Rel.	3 I
	3 I

us-09-724-296-39.rsp

~

```
Bacteria; Proteobacteria; alpha subdivision; Rhizobiaceae group;
                                                                     SEQUENCE FROM N.A.
               Pseudomonas sp.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     NODM_RHILV
                                                                                                                                                                                                                                                                                                                                                                                                                                                                               ACT_SITE
SEQUENCE
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   Query Match
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  P08633;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     NODM_RHILLV
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              Matches
    q
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          Op
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     9
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          ŏ
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              ŏ
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  δ
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       ŏ
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation-the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (see http://www.isb-sib.ch/announce/or send an email to license@isb-sib.ch).
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           13;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            121 VRESSVRAMSAHARVMDGLGLARTPWNLLLLHGGKGGRGAELAALIPDLPDPVRLRLGLE 180
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    76 DDTGA---AVLTHLAPQLLEA-----GHAFTDAGVRLLMHPEQFIVLNSDRPE 120
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                91 RETAVPHDRVSRHIA-SLIETTVYKLDGVDVKFEHVYMQSG-------PADRYE 136
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       137 --STAAHKIAALKTALLGVDCARPSQNLQL-----GSDAVLA-----RVRLELE 178
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 181 NDERAYSPAELLPICEATGTPLVFDAHHHVV----HDKLPDQEDPSVREWVLRARATW-- 234
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          71; Gaps
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            33 RCAAPAERYTDLYD--ANNVRTRQTADSAVSVHKTNLRDERFVHWLRSSNALVPLVRREN 90
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    24 RALSPAEREAKLLDLYSSNIKTLRGAADYCAAHDIRL-----YRLSSSLFPMLDLAG 75
248 SS----PKSEQAIRSHADYVDANFLLPLLERFRQWGTNIDFMIEAKQKDKAL--LRLM 299
                                                                                                                                                                                                                                                                                   "The sequence of the Orgyia pseudotsugata multinucleocapsid nuclear
                                                                                                                                                                                                                                                                                                polyhedrosis virus genome.";
Virology 229:381-399(1997).
-!- FUNCTION: REQUIRED FOR LATE AND VERY LATE GENE EXPRESSION (BY
                                                                                                                                                                 Orgyia pseudotsugata multicapsid polyhedrosis virus (OpMNPV).
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                y match 5.9%; Score 90.5; DB 1; Length 457; Local Similarity 23.1%; Pred. No. 3.2; hes 59; Conservative 34. ....
                                                                                                                                                                                                                                              MEDLINE-97271300; PubMed-9126251;
Ahrens C.H., Russell R.R., Funk C.J., Evans J., Harwood S.,
Rohrmann G.F.;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         Transcription regulation. SEQUENCE 457 AA; 51151 MW; E21E06E50BAC8390 CRC64;
                                                                                                                                                                              dsDNA viruses, no RNA stage; Baculoviridae;
                                                                                         01-NOV-1997 (Rel. 35, Created)
01-NOV-1997 (Rel. 35, Last sequence update)
01-NOV-1997 (Rel. 35, Last annotation update)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          01-0CT-1993 (Rel. 27, Created)
01-0CT-1993 (Rel. 27, Last sequence update)
01-FEB-1996 (Rel. 33, Last annotation update)
PROBABLE ALDEHYDE DEHYDROGENASE (EC 1.2.1.3).
                                                                     457 AA
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      488 AA.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      PRT;
                                                                     PRT;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                             EMBL; U75930; AAC59090.1; -.
                                                                                                                                      LATE EXPRESSION FACTOR 4.
                                                                   STANDARD;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               237 QAPD----STGVK 245
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    STANDARD;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      235 QPPEWQVVHLSNGIE 249
                                                                                                                                                                                          Nucleopolyhedrovirus.
                                                                                                                                                                                                        NCBI_TaxID=164623;
                                                                                                                                                                                                                                 SEQUENCE FROM N.A.
                                                                                                                                                                                                                                                                                                                                          SIMILARITY).
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   DHAL_PSESP
P33008;
                                                                  LEF4_NPVOP
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              Query Match
                                                                                                                                                                              Viruses;
                                        RESULT 2
LEF4_NPVOP
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         RESULT 3
DHAL_PSESP
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          Matches
qq
                                                                                g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        δy
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                op
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  à
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    TO DA
```

```
This SWISS-PROT entry is copyright. It is produced through a collaboration between the SWISS Institute of Bioinformatics and the BMBL outstation—the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (see http://www.isb-sib.ch/announce/or send an email to license@isb-sib.ch).
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      10;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   54 AAHDIRLYRLSSSLFPMLDLAGDDTGAAVLTHLAPQLLEA--GHAFTDAGV----- 102
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                152 HGG-KGGRGAELAALIPDLPDPVRLRLGLENDERAYSPAELLPICEATGTPLVFDAHHHV 210
                                                                                                                       103 -------RILMHPEQFIVLNSDRPEVRESSVRAMSAHARVMDGLGLARTPWNLLLL 151
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  304 IHESIYEAFRDKLVAYAQNVVIGDGSQPGVTMGPLQNAKQFAKVMSLIDAVRQRGGRLIE 363
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      Gaps
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       01-AUG-1988 (Rel. 08, Created)
01-JUN-1994 (Rel. 29, Last sequence update)
20-AUG-2001 (Rel. 40, Last amontation update)
GLUGOSAMINE-FRUCTOSE-6-PHOSPHATE AMINOTRANSFERASE [ISOMERIZING]
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      58;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  NAD (ADP PART) (BY SIMILARITY).
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 5.9%; Score 90; DB 1; Length 488;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 82; Indels
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       BY SIMILARITY.

BY SIMILARITY.

2144AE1EB173C854 CRC64;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 211 VHDKLPDQEDPSVREWVLRARATWQPPEWQVVHLS-NGIEGPQDR 254
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 (EC 2.6.1.16) (GFAT) (NODULATION PROTEIN M).
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                607 AA
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              32; Mismatches
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           Pfam; PF00171; aldedh; 1.
PROSITE; PS00070; ALDEHYDE_DEHYDR_CYS; 1.
PROSITE; PS00687; ALDEHYDE_DEHYDR_GLU; 1.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              Rhizobium leguminosarum (biovar viciae).
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            Pred. No.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      HSSP; P20000; 1A4Z.
InterPro; IPR002086; Aldehyde_dehydr.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                PRT;
                                                                                                       MEDLINE=92332528; PubMed=1629218;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               296 296 483 AA; 52156 MW;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         EMBL; M91440; AAA25995.1; -.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      23.6%;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 53; Conservative
Bacteria; Proteobacteria.
NCBI_TaxID=306;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                STANDARD;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 PIR; S27652; S27652.
PIR; C42971; C42971.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        Oxidoreductase; NAD.
NP_BIND 240 2
ACT_SITE 252 2
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         Best Local Similarity
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           Plasmid sym pRL1JI.
```

```
This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstationher Burden and collaboration its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See http://www.isb-sib.ch/announce/
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  13;
                                                                                                                                  HAIR DEFORMATION
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  Gaps
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 318 IVEIDVASEFRYREIPLS-----PRSAALFISQSGETADTLASLRYCKAHGLRIGAV 369
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              -----LYRLSSSLFPMLDLAGDDTGAA-----VLTHLAPQLLEAGHA---FTDAGV 102
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           103 RLLMHPEQFI-----VLNSDRPEVRESSVRAMSAHARVMDGLGLARTPWNLLLLHGGKG 156
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      SIMILARITY: IN THE C-TERMINAL SECTION; BELONGS TO THE SIS FAMILY. GLMS SUBFAMILY.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      GATASE (BY SIMILARITY).
ISOMBRISATION FRU-6P (BY SIMILARITY).
GLUTAMINE AMIDOTRANSERASE.
7865493083E0A746 CRC64;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         4 LVCLTVGPEVRFRTVTLSRYRALSPAEREAKLLDLYSSNIKTLRGAADYCAAHDIR----
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     157 GRGAELAALIPDLPDPVRLRLGLENDERAYSPAELLPICEATGTPL-VFDAHHHVV----
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  70;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      5.9%; Score 90; DB 1; Length 607; 24.4%; Pred. No. 5.1;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  PROSITE; PS00443; GATASE_TYPE_II; 1.
Transferase; Aminotransferase; Glutamine amidotransferase;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  91; Indels
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               34; Mismatches
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           BY SIMILARITY.
                                                                                                                                                                                                                                                                                                                                                  or send an email to license@isb-sib.ch).
                                                                 MEDLINE=88246045; PubMed=3132583;
                                                                                                                                                                                                                                                                                                                                                                                                                             InterPro; IPR000583; GATase_2.
InterPro; IPR001347; SIS.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                65663 MW;
                                                                                                                                                                                                                                                                                                                                                                           EMBL; Y00548; CAA68626.1; -. PIR; S01040; S01040.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         212 HDKLPDQEDPSVREWVLR 229
                                                                                                                                                                                                                                                                                                                                                                                                                                                         Pfam; PF00310; GATase_2; 1.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   520 HDRFFDKTVSNMQEVAAR 537
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                Conservative
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   602
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           0
                                                                                                                                                                                                                                                                                                                                                                                                                                                                         SIS;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               Plasmid; Nodulation.
                                                                                                                                                                                                                                                                                                                                                                                                      HSSP; P17169; 1MOR.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    Query Match
Best Local Similarity
Matches 63; Conserv
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             607 AA;
                                                                                                                                                                                                               GATASE DOMAIN.
                                      SEQUENCE FROM N.A.
                                                                                                                                                                                                                                                                                                                                                                                                                  MEROPS; C44.971;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      602
           NCBI_TaxID=387;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                      PF01380;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   ACT_SITE
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               SEQUENCE
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           INIT_MET
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              9
Db
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     ò
```

RESULT 5 PUAC_STRLP

```
entities requires a license agreement (See http://www.isb-sib.ch/announce/or send an email to license@isb-sib.ch).
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            This SWISS-PROT entry is copyright. It is produced through a collaboration
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                EMBL outstation
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            There are no restrictions on ng as its content is in no
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             88 PQLLEAGHAFTDAGVRLLMHPEQFIVLNSDRPEVRESSVRAMSAHARVMDGLGLARTP-- 145
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      ------PRMAELSGSRLAAQQQ-MEGLLAPHRPKE 120
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           20-AUG-2001 (Rel. 40, Last annotation update)
NUCLEAR RECEPTOR CO REPRESSOR 2 (N-COR2) (SILENCING MEDIATOR OF
RETINOIC ACID AND THYROID HORMONE RECEPTOR) (SMRT) (SMRTE) (THYROID-,
RETINOIC-ACID-RECEPTOR-ASSOCIATED CO-REPRESSOR) (T3 RECEPTOR-
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                121 PAWFLATVGVSPDHQGK---GLGSAVVLP------GVEAAERAGVPAFLETSAPRN 167
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     74; Gaps
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     43 IKTLRGA-ADYCAA-----HDIRLYRLSSSLFPMLDL-----AGDDTGAAVLTHLA 87
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  20 VRTLAAAFADYPATRHTVDPDRHIERVTELQELFLTRVGLDIGKVWVADD-GAAVAVWTT 78
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             146 --WNLLLL----HGGKGGRGAELAALIPDLPDPVRLRLGLENDERAYSPAEL----
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Musinae; Musi
                                                                                                                                                                                                                                                                          Bacteria; Firmicutes; Actinobacteria; Actinobacteridae;
Actinomycetales; Streptomycineae; Streptomyces.
                                                                                                                                                                                                                                                                                                                                                                                                                                                         Lacalle R.A., Pulido D., Vara J., Zalacain M., Jimenez A.: "Molecular analysis of the pac gene encoding a puromycin N-acetyl transferase from Streptomyces alboniger."; Gene 79:375-380(1989).
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  Mus musculus (Mouse).
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  66; Indels
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  Length 199;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         between the Swiss Institute of Bioinformatics and the
the European Bioinformatics Institute. There are no re
use by non-profit institutions as long as its conten
modified and this statement is not removed. Usage by
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    Antibiotic resistance; Transferase; Acyltransferase.
SEQUENCE 199 AA; 21496 MW; BEDA36E20304DA35 CRC64;
                                                                                                                                                                                                                                       Streptomyces lipmanii (Streptomyces alboniger).
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           5.8%; Score 89; DB 1; 26.3%; Pred. No. 1.5;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    192 LPICEATGTPLVFDAHHHVVHDKLPDQEDPSVREWVL 228
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                168 LPFYERLGFTVTADV-----EVP--EGP--RTWCM 193
                                                                                                   01-FEB-1996 (Rel. 33, Last sequence update) 20-AUG-2001 (Rel. 40, Last annotation update) PUROMYCIN N-ACETYLTRANSFERASE (EC 2.3.-.-).
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             20-AUG-2001 (Rel. 40, Created)
20-AUG-2001 (Rel. 40, Created)
20-AUG-2001 (Rel. 40, Last sequence update)
20-AUG-2001 (Rel. 40, Last annotation update
   199 AA
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           PRT; 2472 AA
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  Mismatches
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     -!- FUNCTION: DETOXIFICATION OF PUROMYCIN.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                InterPro; IPR000182; Acetyltransf_GCN5
Pfam; PF00583; Acetyltransf; 1.
                                                                                                                                                                                                                                                                                                                                                                                                                                                 MEDLINE=90006767; PubMed=2676728;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              20;
                                                                01-JAN-1990 (Rel. 13, Created)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    EMBL; M25346; AAA64928.1; -.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               ASSOCIATING FACTOR) (TRAC). NCOR2 OR SMRT.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              57; Conservative
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           STANDARD;
   STANDARD;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         79 PESVEAGAVFAEIG----
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              Query Match
Best Local Similarity
                                                                                                                                                                                                                                                                                                                                                                                 [1]
SEQUENCE FROM N.A.
                                                                                                                                                                                                                                                                                                                                       NCBI_TaxID=1892;
PUAC_STRLP
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           NCR2_MOUSE
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    NCR2_MOUSE
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  Matches
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         qq
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           οy
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         q
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             δ
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             Οy
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                Пp
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                õ
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                q
```

```
This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See http://www.isb-sib.ch/announce/or send an email to license@isb-sib.ch).
                                                                                                                                                                                                                                                                                                                                                                   SUBUNIT: FORMS A LARGE COREPRESOR COMPLEX THAT CONTAINS SIN3A/B AND HISTORE DEACETYLASES HDACI AND HDACZ. THIS COMPLEX ASSOCIATES WITH THE THYROID (TR) AND THE RETINOID ACID RECEPTORS (RAR) IN THE ABSENCE OF LIGAND, AND MAY STABILIZE THEIR INTERACTION WITH TFIIB. SUBCELLULAR LOCATION: NUCLEAR.

ALTERNATIVE PRODUCTS: 2 ISOFORMS; ALPHA (SHOWN HERE) AND BETA; ARE PRODUCED BY ALTERNATIVE SPLICING.

TISSUE SPECIFICITY: UBIQUITOUS. ALSO WIDELY EXPRESSED IN EARLY
                                                                                                                                                                                                                                                                                          Proc. Natl. Acad. Sci. U.S.A. 96:3519-3524(1999).
-1- FUNCTION: MEDIATES THE TRANSCRIPTIONAL REPRESSION ACTIVITY OF SOME NUCLEAR RECEPTORS BY PROMOTING CHROMATIN CONDENSATION, THUS PREVENTING ACCESS OF THE BASAL TRANSCRIPTION.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    INTERACTION WITH SIN3A/B (BY SIMILARITY).
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              DOMAINS (ID1 AND ID2).

DOMAINS THE TWO INTERACTION DOMAINS (ID) CONTAIN A CONSERVED SEQUENCE REFERRED TO AS THE CORRE BOX. THIS MOTIF IS REQUIRED AND SUFFICIENT TO PERMIT BINDING TO UNLIGANDED TR AND RARS. SEQUENCES
                               TISSUE=Spleen, and Brain;
MEDLINE-99178941; PubMed-10077563;
Ordentlich P., Downes M., Xie W., Genin A., Spinner N.B., Evans R.M.;
"Unique forms of human and mouse nuclear receptor corepressor SMRT.";
Proc. Natl. Acad. Sci. U.S.A. 96:2639-2644(1999).
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 DOMAIN: THE N-TERMINAL REGION CONTAINS REPRESSION FUNCTIONS THAT ARE DIVIDED INTO THREE INDEPENDANT REPRESSION DOMAINS (RDI, RDZ AND RD3). THE C-TERMINAL REGION CONTAINS THE NUCLEAR RECEPTOR-INTERACTING DOMAINS THAT ARE DIVIDED IN TWO SEPARATE INTERACTION
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    SIMILARITY: CONTAINS 1 SANT DOMAIN (COMPOSED OF SANT-A+SANT-B). SIMILARITY: CONTAINS 2 CORNR BOX. SIMILARITY: BELONGS TO THE N-COR NUCLEAR RECEPTOR COREPRESSORS
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              SMARY; SM00395; SANT; 2.
PROSITE; PS50090; MYB_3; 1.
Nuclear protein; Transcription regulation; DNA-binding; Repressor;
                                                                                                                                                                                                               Park E.J., Schroen D.J., Yang M., Li H., Li L., Chen J.D.; "SMRTe, a silencing mediator for retinoid and thyroid hormone receptors-extended isoform that is more related to the nuclear
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              FLANKING THE CORNR BOX DETERMINE NUCLEAR HORMONE RECEPTOR
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       SANT-A (POTENTIAL).
SANT-B (POTENTIAL).
COILED COIL (POTENTIAL).
COILED COIL (POTENTIAL).
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 COILED COIL (POTENTIAL)
                SEQUENCE FROM N.A. (ISOFORM ALPHA AND ISOFORM BETA).
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             CORNR BOX OF ID1. CORNR BOX OF ID2.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         PRO-RICH.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         PRO-RICH.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            PRO-RICH.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     POLY-GLN.
                                                                                                                                                      SEQUENCE FROM N.A. (ISOFORM ALPHA).
                                                                                                                                                                           TISSUE-Embryo;
MEDLINE-99199215; PubMed-10097068;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       MGD; MGI:1337080; Ncor2.
Interpro; IPR001005; Myb_DNA_bind.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       Pfam; PF00249; myb_DNA-binding; 2.
SMART; SM00395; SANT; 2.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               Coiled coil; Alternative splicing
DOMAIN 165 207 COIL
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             EMBL; AF113001; AAD20944.1; -.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       EMBL; AF125671; AAD22972.1; -.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    AF113002; AAD20945.1; -.
                                                                                                                                                                                                                                                                         receptor corepressor.";
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          653
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          1357
2098
2300
507
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 SPECIFICITY.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       429
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  EMBRYOS.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              FAMILY.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         DNA_BIND
DNA_BIND
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    DOMAIN
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            DOMAIN
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            DOMAIN
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               DOMAIN
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  DOMAIN
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     DOMAIN
```

```
HHLPHPRLLWTRMNKKPRLLQLPRQRMPRSRSLRPRRSMWE
                                                                                                                                                                                                                                                                     GSATSGSITKGLPSTRADGPSYRGSITHG -> APPPVEA
                                                                                                      -> PSPAAPPATVDKFFQEAPAAPAQTEDAKEQKSEAEEI
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               ---HLSNGIEG 250
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   815 WIRMNKKPRLLOLPRORMPRSRS--LRPRRSMWEKPEEPEASEEPPESVKSDHKEETEEE 872
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 157 GRGAELAALIPDLPDPVRLRLGLENDERAYSPAELLPICEATGTPLVFDAHHHVVHDK-- 214
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      765 GTEALPAATQPPVPPP-----EEPAVAPAEPSPVPDASGPPSP-EPSHHLPHPRLL 814
                                                                                                                                                                                                                                                                                                                                                                                                                                                                           42; Gaps
                                                                                                                           E -> K (IN REF. 1; AAD20945).
E -> K (IN REF. 1; AAD20945).
E -> K (IN REF. 1; AAD20945).
E -> K (IN REF. 1; AAD20945).
S -> F (IN REF. 1; AAD20944).
I -> IQ (IN REF. 1) AAD20944).
PKLPTEPPRWSSGLPFPI -> QSYRLSPHAGHRLPSH
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              MEDLINE-95273376; PubMed=7753822;
Helaakoski T., Annunen P., Vuori K., Macneil I.A., Pihlajaniemi T.,
Kivirikko K.I.;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus
                                                                                                                                                                                                                                                                                   SPRASQYPGCRRPQLQRLYHPR (IN REF. 2).
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 01-NOV-1997 (Rel. 35, Last sequence update)
30-MAY-2000 (Rel. 39, Last amnotation update)
PRODYL 4-HYDROXXLASE ALPHA-2 SUBUNIT PRECURSOR (EC 1.14.11.2).
                                   M \rightarrow RL (IN REF. 2). PPMLYDA \rightarrow RHVVRR (IN REF. 2).
                                                                                                                                                                                                                         TRADPL (IN REF. 2)
                                                                                                                                                                                                                                                                                                                                                                                                                                                   5.8%; Score 89; DB 1; Length 2472; 25.3%; Pred. No. 35;
                                                        D -> H (IN REF. 1; AAD20944).
T -> M (IN REF. 1; AAD20944).
V -> A (IN REF. 2).
                                                                                                                                                                                                                                                                                                                                                                                                                                                                         20; Mismatches 53; Indels
                                                                                                                                                                                                                                                                                                                                                                                                     -> A (IN REF. 2).
: 2A58F4DF7B79285B CRC64;
                       MISSING (IN ISOFORM BETA).
                                                                                                                                                                                                                                                                                                                                                                             TA -> AV (IN REF. 2).
LE -> SK (IN REF. 2).
T -> A (IN REF. 2).
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               ---LPDQEDPSVREWVLR-ARATWQPPEWQVV---
                                                                                                                                                                                                                                                                                                                  A -> P (IN REF. 2).
A -> G (IN REF. 2).
C -> A (IN REF. 2).
C -> S (IN REF. 2).
A -> G (IN REF. 2).
                                                                                                                                                                                                                                    MISSING (IN REF. 2).
MISSING (IN REF. 2).
                                                                                                                                                                                                                                                           -> E (IN REF. 2)
                                                                                                                                                                                                                                                                                             A -> S (IN REF. 2
MISSING (IN REF.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               PRT; 537 AA.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              251 PQDRRHSHLIADFPSAYADVPQIEVEAKGKEEAI 284
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           DVG (IN REF.
                                                                                                                                                                                                              (IN REF. 2). PHAADPSA ->
                                                                                                                                                                                                                                                                                                                                                                                                                AA; 270856 MW;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      01-Nov-1997 (Rel. 35, Created)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                         39; Conservative
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              STANDARD;
1619
254
254
176
402
555
756
785
846
                                                                                                                                                   867
895
916
975
1063
                                                                                                                                                                                                                                    1133
1149
1157
1201
                                                                                                                                                                                                                                                                                                                               1913
1923
1956
1968
2196
2214
2224
                                                                                                                                                                                                                                                                                                                     606
                                                                                                                                                                                                                                                                                                         1857
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               Mus musculus (Mouse).
                                                                                                                                                                                                                                                                                                                                                                                                                                                   Query Match
Best Local Similarity
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       SEQUENCE FROM N.A.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                NCBI_TaxID=10090;
                                                                                                                                                              895
916
975
                                                                                                                                                                                                                                      1133
                                                                                                                                                                                                                                                                                                                                1913
1923
1956
1968
2195
2213
2224
2472
                                                                                                                                                                                                 1046
                                                                                                                                                                                                                         1073
                                                                                                                                                                                                                                                                                             9691
                                                                                                                                                                                                                                                                                                         1855
                                                                                                                                                                                                                                                                                                                     1909
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           P4H2_MOUSE
Q60716;
                                                                                                                                                              CONFLICT
CONFLICT
CONFLICT
CONFLICT
                                                                   CONFLICT
                                                                                                                                                                                                                                    CONFLICT
                                                                                                                                                                                                                                                                                                                               CONFLICT
                                                                                                                                                                                                                                                                                                                                                      CONFLICT
                                 CONFLICT
                                                                                                                                                                                                                                                                                                                                                                                                                 SEQUENCE
                       VARSPLIC
                                                         CONFLICT
                                                                                                                                                                                                                         CONFLICT
                                                                                                                                                                                                                                                                      CONFLICT
                                                                                                                                                                                                                                                                                                           CONFLICT
                                                                                                                                                                                                                                                                                                                     CONFLICT
                                                                                                                                                                                                                                                                                                                                                                              CONFLICT
                                                                                                                                                                                                                                                                                                                                                                                         CONFLICT
                                                                                                                                                                                                                                                                                                                                                                                                      CONFLICT
                                                                                           CONFLICT
                                                                                                                              CONFLICT
                                                                                                                                          CONFLICT
                                                                                                                                                      CONFLICT
                                                                                                                                                                                                                                                           CONFLICT
                                                                                                                                                                                                                                                                                             CONFLICT
           NOMAIN
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     P4HA2.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   P4H2_MOUSE
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               215
                                                                                                                                                                                                                                                                                                                                                                                                                                                                           Matches
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      RESULT
 qq
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               Qγ
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      qq
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               õ
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               П
```

ω ω

```
This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See http://www.isb-sib.ch/announce/or send an email to license@isb-sib.ch).
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    22;
                                                                                                                                                                                   -!- COFACTOR: IRON AND ASCORBATE.
-!- COFACTOR: IRON AND ASCORBATE.
-!- ENZYME REGULATION: IN CONTRAST TO THE ALPHA-1 SUBUNIT, ALPHA-2 IS
INHIBITED BY POLY(L-PROLINE) ONLY AT VERY HIGH CONCENTRATIONS.
-!- SUBUNIT: TETRAMER OF TWO ALPHA CHAINS AND TWO BETA CHAINS (THE
BETA CHAIN IS THE MULTI-FUNCTIONAL PDI).
-!- SUBCELLULAR LOCATION: ENDOPLASMIC RETICCILUM LUMEN.
-!- TISSUE SPECIFICITY: EXPRESSED IN A VARIETY OF TISSUES.
                                                                               Proc. Natl. Acad. Sci. U.S.A. 92:4427-4431(1995).
-!- FUNCTION: CATALYZES THE POSTTRANSLATIONAL FORMATION OF 4-
HYDROXYPROLINE IN COLLAGENS.
-!- CATALYTIC ACTIVITY: PROCOLLAGEN L-PROLINE + 2-OXOGUJTARATE + O(2)
- PROCOLLAGEN TRANS-4-HYDROXY-L-PROLINE + SUCCINATE + CO(2).
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     343 YYDVMSDEEIERIKEIAKPKLARATVRDPKTGVLTVASYRVSKSSWLEEDDDPVVARVNR 402
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                54 EAKLAKIKSWASKMEALTSRSAADPEGY-LAHPVNAYKLVKRLNTDWPALGDLVLQDASA 112
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        81 AVLTHLAPQLL------EAGHAFTDAGVRLLMHPEQFIVLNSD---RPEVRESSVRAMS 130
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  ---RGAELAALIPDLPOPVRLRLG-----LENDERAYS------P 188
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        167 S---VDDCFGLGRSAYNEGDYYHTVLWMEQVLKQLDAGEEATVTKSLVLDYLSYAVFQLG 223
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     283 TDYLPERDVYESLCRGEGVKLTPRROKKLFCRYHHGNRVPQLLIAPFKEEDEWDSPHIVR 342
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  212 -HDKLPDQEDPSVREWV--LRARATWQPPEWQV-----VHLSNGIEGPQD-----R 254
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                39; Mismatches 116; Indels 160; Gaps
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        32 EAKLLDLYSSNIK----TLRGAAD---YCAAHDIRLYRLSSSL---FPML-DLAGDDTGA 80
"Cloning, baculovirus expression, and characterization of a second mouse prolyl 4-hydroxylase alpha-subunit isoform: formation of an alpha 2 beta 2 tetramer with the protein disulfide-isomerase/beta
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            PROLYL 4-HYDROXYLASE ALPHA-2 SUBUNIT.
N-LINKED (GLCNAC. . .) (POTENTIAL).
N-LINKED (GLCNAC. . .) (POTENTIAL).
C470976C75E82518 CRC64;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      Pfam; PF02715; P4-hydrxy_alpha; 1.
Oxidoreductase: Dioxygenase; Iron; Vitamin C; Glycoprotein;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      5.8%; Score 88; DB 1; Length 537; 22.2%; Pred. No. 6.4;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    255 RHSHL-----IADFPSAYADVPQIEVEAKGKEEAIAAL 287
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                | | : | : | : | | RMOHITGLTVKTAELLQVANYGMGGQYEPHFDFSRSDDEDAFKRL 447
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                189 AELLP-----ICEATGTPL-----VFDAHH------
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            131 AHARVMDGLGLARTPWN---------LLLLHGGKGG---
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        POTENTIAL.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      Endoplasmic reticulum; Signal.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        537 AA; 61002 MW;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 EMBL; U16163; AAC52198.1; -.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                90; Conservative
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            Similarity
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            CARBOHYD
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      Query Match
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    CARBOHYD
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        SEQUENCE
                                                                  subunit.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              Sest Local
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      SIGNAL
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              Matches
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              403
  ö
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                q
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             QQ
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 Q
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            õ
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                à
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     q
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          9
```

```
This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation-the Buropean Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed 'Usage by and for commercial entities requires a license agreement (See http://www.isb-sib.ch/announce/
                                                                                                                                                                                                                                                                                   MEDLINE-9825987; PubMed=9634230; Garnier T., Churcher C., Harris D., Cole S.T., Brosch R., Parkhill J., Garnier T., Churcher C., Harris D., Gordon S.V., Eiglumeier K., Gas S., Barry C.E. Ill, Trekaia F., Badcock K., Basham D., Brown D., Chillingworth T., Connor R., Davies R., Devlin K., Feltwell T., Gentles S., Hamlin N., Holroyd S., Hornsby T., Jagels K., Krogh A., McLean J., Moule S., Murphy L., Oliver S., Seeger K., Skelton S., Squares S., Murphy L., Rutter S., Seeger K., Skelton S., Squares R., Sulston J.E., Taylor K., Whitehead S., Barrell B.G.; "Deciphering the biology of Mycobacterium tuberculosis from the complete genome sequence."
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      Fleischmann R.D., Alland D., Eisen J.A., Carpenter L., White O., Peterson J., DeBoy R., Dodson R., Gwinn M.L., Haft D., Hickey E., Kolonay J.F., Nelson W.C., Umayam L.A., Ermolaeva M.D., Salzberg S.L., Delcher A., Utterback T., Weidman J., Khouri H., Gill J., Mikula A.,
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           63 L.S.S.S.L.F.PMLDI.AGDDTGAAVLTHLAPQLI.EAGHAFTDAGVRLLMHPEQFIVLNSDRPEV- 121
                                                                                                                                                               Bacteria; Firmicutes; Actinobacteria; Actinobacteridae;
Actinomycetales; Corynebacterineae; Mycobacteriaceae; Mycobacterium.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    86; Indels 94; Gaps
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             "Whole genome comparison of Mycobacterium tuberculosis clinical and
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        PROSITE; PS00107, PROTEIN_KINASE_ATP; 1.
PROSITE; PS00108; PROTEIN_KINASE_ST; FALSE_NEG.
PROSITE; PS50011; PROTEIN_KINASE_DOM; 1.
Hypothetical protein; Transferase; Serine/threonine-protein kinase;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            laboratory strains.";
Submitted (APR-2001) to the EMBL/GenBank/DDBJ databases.
-!- SIMILARITY: BELONGS TO THE SER/THR FAMILY OF PROTEIN KINASES.
                                                            20-AUG-2001 (Rel. 40, Last sequence update)
20-AUG-2001 (Rel. 40, Last annoctation update)
PROBABLE SERINETHREONINE-PROTEIN KINASE PKNK (EC 2.7.1.-)
PKNK OR RV3080C OR MT3165 OR MTV013.01C OR MTCY22D7.01.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           Ouery Match 5.6%; Score 86; DB 1; Length 1110; Best Local Similarity 21.8%; Pred. No. 23;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 2F81BF18BCD3E472 CRC64;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  ATP (BY SIMILARITY).
ATP (BY SIMILARITY).
      PRT; 11110 AA
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               PROTEIN KINASE
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      Mismatches
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         BY SIMILARITY
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             InterPro; IPR000719; Euk_pkinase.
InterPro; IPR002290; Ser_thr_kin_actsite.
Pfam; PF00069; pkinase; 1.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     or send an email to license@isb-sib.ch).
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         148 148 BY
1110 AA; 119417 MW;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    28;
                                        20-AUG-2001 (Rel. 40, Created)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       ATP-binding; Complete proteome.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 EMBL; AE007133; AAK47501.1; -.
TIGR; MT3165; -
                                                                                                                                                Mycobacterium tuberculosis.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               SEQUENCE FROM N.A.
STRAIN=CDC 1551 / Oshkosh;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                EMBL; Z83866; CAB06268.1;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        Nature 393:537-544(1998).
    STANDARD;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      58; Conservative
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                40
55
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          TubercuList; Rv3080c;
                                                                                                                                                                                                                                                 SEQUENCE FROM N.A.
                                                                                                                                                                                                            NCBI_TaxID=1773;
                                                                                                                                                                                                                                                                         STRAIN=H37RV:
PKNK_MYCTU
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       Bishai W.;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         ACT_SITE
SEQUENCE
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  NP BIND
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      BINDING
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               DOMAIN
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      Matches
    õ
```

72751 MW; 56516DC3EA40E4B0 CRC64;

653 AA;

```
SEQUENCE
                                                                                                                                                                                                                                                                                                                                                                                                                                                      δλ
                                                                                                                                                                                                                                 G
D
                                                                                                                                                                                                                                                                                    qq
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           between the Swiss Institute of Bioinformatics and the EMBL outstation the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See http://www.isb-sib.ch/announce/or send an email to license@isb-sib.ch).
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    This SWISS-PROT entry is copyright. It is produced through a collaboration
 --NDDRIAVV 459
                           122 ----RESSVRAMSAHARVMDG----LGLARTPWNLLLLHGGKGGRGAELAALIPDLPDP 172
                                                                                                           515 AALR--FDTDEAAALLNDAGGLRLPRADVQALTTSTDGWAAALRLAALSLRGGGDATQLL 572
                                                                                                                                     199 -GTPLVFDAHHHVVHDKLPDQEDPSVREWVLRA--------RATWQPP 237
                                                                                                                                                               573 RCLSGASDVIHEFLSENVLDTLEPELREFLLVASVTERTCGGLASALAGITNGRAMLEEA 632
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   "Identification of a mouse brain cDNA that encodes a protein related to the Alzheimer disease-associated amyloid beta protein precursor.";
Proc. Natl. Acad. Sci. U.S.A. 89:10758-10762(1992).
-!- SUBCELLULAR LOCATION: TYPE I MEMBRANE PROTEIN. MAY BE PROCESSED IN THE GOLGI COMPLEX; HOWEVER, IT IS NOT YET CLEAR WHETER APLP IS SECRETED.
                                                                                                                                                                                                                                                                                                                                                                                                  Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
                                                    460 IDDWHRVSDSRTQAALGFLLDNGCHHLQLIVTSWSRAGLPVGRLRIGDELAEI----DS
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            N-LINKED (GLCNAC. . .) (POTENTIAL).
N-LINKED (GLCNAC. . .) (POTENTIAL).
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            Wasco W., Bupp K., Magendantz M., Gusella J.F., Tanzi R.E.,
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    CLATHRIN-BINDING (POTENTIAL).
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           EXTRACELLULAR (POTENTIAL).
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     CYTOPLASMIC (POTENTIAL)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                AMYLOID-LIKE PROTEIN 1
                                                                                 173 VRLRLGLENDERA-----YSPAELLPICEAT----
                                                                                                                                                                                                                                                                                                               01-0cT-1993 (Rel. 27, Created)
01-0cT-1993 (Rel. 27, Last sequence update)
01-UUN-1994 (Rel. 29, Last annotation update)
01-UUN-1994 (Rel. 29, Last annotation update)
                                                                                                                                                                                                                                                                                        653 AA
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       -! - SIMILARITY: BELONGS TO THE APP FAMILY.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       Signal.
POTENTIAL.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           POTENTIAL
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  POLY-GLU.
422 LAESLGHVLEEHGDDAGRYVLTSLIDEIHE-
                                                                                                                                                                                          238 EWQVVHLSNGIEGPQDRRHSHLIADF 263
                                                                                                                                                                                                                    633 EHRGLFLORTEDDPNWFRFHOMFADF 658
                                                                                                                                                                                                                                                                                        PRT;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  MEDLINE=93066322; Pubmed=1279693;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            SMART; SM00006; A4_EXTRA; 1.
PROSITE; PS00319; A4_EXTRA; 1.
PROSITE; PS00320; A4_INTRA; 1.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              Glycoprotein; Transmembrane; SIGNAL
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             EMBL; L04538; AAA37247.1; -.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       MGD; MGI:88046; Aplp1.
Interpro; IPR001886; A4_APP.
Pfam; PF02177; A4_EXTRA; 1.
PRINTS; PR00203; AMYLOIDA4.
                                                                                                                                                                                                                                                                                        STANDARD;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              653
6606
653
646
271
464
                                                                                                                                                                                                                                                                                                                                                                                     Mus musculus (Mouse).
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            A46362
                                                                                                                                                                                                                                                                                                                                                                                                                                                     SEQUENCE FROM N.A.
                                                                                                                                                                                                                                                                                                                                                                                                                               NCBI_TaxID=10090;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                      TISSUE-Brain;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             PIR; A46362;
                                                                                                                                                                                                                                                                                       APP1_MOUSE
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             Solomon F.;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           PRANSMEM
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               CARBOHYD
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            CARBOHYD
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             DOMAIN
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       DOMAIN
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          DOMAIN
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  DOMAIN
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 CHAIN
                                                                                                                                                                                                                                                                          APP1 MOUSE
                                                                                                                                                                                                                                                                                                                                                                        APLP1
                                                                                                                                                                                                                                                                                                    q
                            Ωý
                                                     g
                                                                                 ò
                                                                                                                                      ōλ
                                                                                                                                                               ΩD
                                                                                                                                                                                          ŏ
```

```
This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See http://www.isb-sib.ch/announce/
                                                                                  10;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            -!- SUBCELLUIAR LOCATION: NUCLEAR (BY SIMILARITY).
-!- SIMILARITY: BELONGS TO THE BASIC HELIX-LOOP-HELIX (BHLH) FAMILY OF
TRANSCRIPTION FACTORS. "HAIRY" SUBFAMILY.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                20-AUG-2001 (Rel. 40, Last sequence update)
20-AUG-2001 (Rel. 40, Last annotation update)
HAIRY/ENHANCER-OF-SPLIT RELATED WITH YRPW MOTIF 1 (HAIRY AND ENHANCER
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 -:- SUBUNIT: EFFICIENT DNA BINDING REQUIRES DIMERIZATION WITH ANOTHER BHLH PROTEIN (BY SIMILARITY).
                                                                                  Gaps
                                                                                                                                                                                                                               83 CLLDPQRVLEYCRQMYPELHIARVEQAAQAIPMERWCGGTRSGRCAHPHHEVVPFHCLPG 142
                                                                                                                                                                                                                                                                                                                                                                                                                                                        148 LLILLHGGKGGRGAELAALLPDLPDPVRLRLGLENDERAYSPAELLPICEATGTPLVFDAH 207
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  208 HHVVHDKLPDOEDPSVREWVLRARATWQPPEWQVVHLSNGIEGPQDRRHSHLIADFPSA- 266
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                95 HAFTDAGVRLLMHPEQFIVLNSDRPEVRESSVR-----AMSAHARVMDGLGLARTPWN 147
                                                                                                                                                                                                                                                                                                                                                                             143 EFVSEA----LLVPEGCRFLHQERMDQCESSTRRHQEAQEACSSQGLILHGSG----- 191
                                                                                                                                                         53 CAAHDIRLYRLSSSLFPMLDLA-------GDDTGAAVLTH--LAPQLLEAG 94
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            Eukaryota, Metazoa, Chordata, Craniata, Vertebrata, Euteleostomi,
Mammalia, Eutheria, Carnivora, Fissipedia, Canidae, Canis.
                                                                          :66
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               "Functional cloning of nuclear proteins and nuclear targeting sequences.";
Query Match 5.6%; Score 85.5; DB 1; Length 653; Best Local Similarity 21.2%; Pred. No. 13; Matches 55; Conservative 22; Mismatches 83; Indels 99
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    Pichon B., Mercan D., Pouillon V., Christophe-Hobertus C.,
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      Submitted (JAN-2000) to the EMBL/GenBank/DDBJ databases.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           Transcription regulation; DNA-binding; Nuclear protein.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              304 AA.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             or send an email to license@isb-sib.ch).
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           PROSITE; PS00038; HELIX_LOOP_HELIX; 1.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              PRT;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 20-AUG-2001 (Rel. 40, Created)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                EMBL; AJ388551; CAB65543.1; -.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           InterPro; IPR003015; HLH_Myc.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           InterPro; IPR001092; HLH_dim.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  SPLIT RELATED-1.) (HESR-1).
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      : : || | : :||
253 DDYFVEPPQABEBEBEBE 271
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              267 ---YADVPOIEVEAKGKEE 282
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              InterPro; IPR003650; Orange.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     SMART; SM00353; HLH; 1.
SMART; SM00511; ORANGE; 1.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              STANDARD;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             Canis familiaris (Doq).
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 Pfam; PF00010; HLH; 1.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             SEQUENCE FROM N.A.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   NCBI_TaxID=9615;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     TISSUE=Thyroid;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 HEY1 OR HESRI
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               Christophe D.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              HEY1_CANFA
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 Q9TSZ2
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          HEY1_CANFA
```

```
13;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       124 -----SLGFRECLAEVARYLS--IIEGLDASDPLRVRLVSH------LNNYASQ-REA-- 167
                                                                                                                                                                                                                                                                                                      127 RAMSAHARVMDGLGLARTPWN------LLLLHGGKGGRG------AELA 163
                                                                                                                                                                                                                                                                                                                                      168 -ASGAHA-----GLGHLPWGSAFGHHPHVAHPLLLPQSGHGNTGTSASPTDPHHQGRLA 220
                                                                                                                                                                                                         68 SLSELRRLVPSAFEKQGSAKLEKAEILQMTVDHLKMLHTAGGKGYFDAHALAMDYR---- 123
                                                                                                                                                                                                                                         67 LFPMLDLAGDDTGAAVLTHLAPQLLEAGHAFTDAGVRLLMHPEQFIVLNSDRPEVRESSV 126
                                                                                                                                            Gaps
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            of the bare
                                                                                                                                                                         19 TLSRYRALSPAERE-----AKLLDLYSSNIKTLR--GAADYCAAHDIRL-YRLSSS 66
             HELIX-LOOP-HELIX MOTIF (BY SIMILARITY).
YRPW MOTIF (REQUIRED FOR ACTIVITY)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            Nagarajan U.M., Louis-Plence P., DeSandro A., Nilsen R., Bushey A.,
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             Nagarajan U.M., Louis-Plence P., DeSandro A., Nilsen R., Bushey A.,
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               Eukaryota, Metazoa, Chordata, Craniata, Vertebrata, Buteleostomi,
Mammalia, Eutheria, Primates, Catarrhini, Hominidae, Homo.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              Steimle V., Durand B., Barras E., Zufferey M., Hadam M.R., Mach B., Reith W.:
A novel DNA-binding regulatory factor is mutated in primary MHC class II deficiency (bare lymphocyte syndrome).";
Genes Dev. 9:1021-1032(1995).
                                                                                                                                                                                                                                                                                                                                                                      164 ALIPDLP-----DPVRLRLGLENDERAYSPAELLPICEATGTPLVFDAHH 208
                                                                                                                                                                                                                                                                                                                                                                                                      221 AAHPEAPALRAPPSGGLGPV---LPVVTSASKLSPPLLSSVASLSAFPFSFGSFH 272
                                                                                                                                          75;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          Villard J., Peretti M., Masternak K., Barras E., Caretti G.,
Mantovani R., Reith W.;
                                                                                                            5.5%; Score 84.5; DB 1; Length 304;
24.7%; Pred. No. 6.1;
tive 23; Mismatches 79; Indels 79
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          PRTX-B is the gene responsible for the most common cause lymphocyte syndrome, an MHC class II immunodeficiency."; Immunity 10:153-162(1999).
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  DNA-BINDING PROTEIN RFX5 (REGULATORY FACTOR X SUBUNIT 5)
                                                           191D5BD42970F412 CRC64;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 01-FEB-1996 (Rel. 33, Created)
01-FEB-1996 (Rel. 33, Last sequence update)
20-AUG-2001 (Rel. 40, Last annotation update)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      616 AA
BASIC DOMAIN.
                                           (POTENTIAL)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      PRT;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               TISSUE-Lymphoblast;
MEDLINE-99170284; PubMed-10072068;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           MEDLINE=20242030; PubMed=10779326;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 SEQUENCE FROM N.A.
MEDLINE=95262896; Pubmed=7744245;
                                                            304 AA; 32471 MW;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          Immunity 10:399-399(1999).
                                                                                                                             Local Similarity 24.7 es 58; Conservative
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      STANDARD;
62
105
297
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   Homo sapiens (Human)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            CHARACTERIZATION.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              PARTIAL SEQUENCE
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 NCBI_TaxID=9606;
50
63
294
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      RFX5_HUMAN
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           Boss J.M.;
DNA_BIND
                                                              SEQUENCE
                                                                                                               Query Match
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            ERRATUM
                DOMAIN
                                DOMAIN
                                                                                                                                              Matches
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         FFF
                                                                                                                                                                                                                                           ò
                                                                                                                                                                                                                                                                         g
                                                                                                                                                                                                                                                                                                         ò
                                                                                                                                                                                                                                                                                                                                        g
                                                                                                                                                                                                                                                                                                                                                                        ò
                                                                                                                                                                                                                                                                                                                                                                                                    g
```

```
This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation-the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            entities requires a license agreement (See http://www.isb-sib.ch/announce/
or send an email to license@isb-sib.ch).
RECOGNIZES X-BOXES, MEDIATES COOPERATIVE BINDING BETWEEN RFX AND NF-Y. REX BINDS THE X1 BOX OF MHC-II PROMOTERS.
SUBUNIT: REX CONSISTS OF AT LEAST THREE DIFFERENT SUBUNITS; RFXAP, REX5 AND RFX-B/FRXANK, WI'LH EACH SUBUNIT REPRESENTING A SEPARATE COMPLEMENTATION GROUP. RFX FORMS COOPERATIVE DNA BINDING COMPLEXES WITH X2BP AND CBF/NF-Y. RFX ASSOCIATES WITH CIITA TO FORM AN ACTIVE TRANSCRIPTIONAL COMPLEX.
                                                                                                                                                                                                                                          SUBCELLULAR LOCATION: NUCLEAR.
TISSUE SPECIFICITY: UBIQUITOUS.
DOMAIN: THE N-TERMINUS IS REQUIRED FOR 1TS ASSOCIATION WITH REXANK
AND REXAP, FOR ASSEMBLY OF THE REX COMPLEX, AND FOR BINDING OF
THIS COMPLEX TO ITS X BOX TRAGET SITE IN THE MHC-1I PROMOTER. THE
C-TERMINUS MEDIATES COOPERATIVE BINDING BETWEEN THE REX COMPLEX
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        47 RGAADYCAAHDIRLYRLSSSLFPMLDLAGDDT---GAAVLTHLAPQLLEAGHAFT-DAGV 102
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 154 RGQSKYCYSGIRRKTLVSMPPLPGLDLKGSESPEMGPEVTPAPRDELVEAACALTCDWAE 213
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 CLASS IT-DEFICIENT COMBINED IMMUNODEFICIENCY); A FORM OF SEVERE COMBINED IMMUNODEFICIENCY); A FORM OF SEVERE COMBINED IMMUNODEFICIENCY DISEASE (SCID). RFX5 IS LINKED WITH BLS COMPLEMENTATION CERTIES.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        PTM: PHOSPHORYLATED.
DISEASE: DEFECTS IN RFX5 ARE A CAUSE OF HEREDITARY MHC CLASS II
DEFICIENCY (ALSO KNOWN AS BARE LYMPHOCYTE SYNDROME (BLS) OR HLA
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              InterPro; IPR003150; RFX_DNA_binding.
Pfam; PF02257; RFX_DNA_binding; 1.
DNA-binding; Transcription regulation; Activator; Nuclear protein;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          RNA POLYMERASE SIGMA FACTOR RPOD (SIGMA-A) (MAJOR VEGETATIVE SIGMA
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     Clostridium acetobutylicum.
Bacteria; Firmicutes; Bacillus/Clostridium group; Clostridiaceae;
Clostridium.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        DB 1; Length 616;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            Indels
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     5EBB33C677BB717F CRC64;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                103 RILMHPEQFIVINSDRPEVRESSVRAMSAHARVMDGLGLA 142
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   48;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          (Rel. 28, Last sequence update) (Rel. 40, Last annotation update)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          378 AA
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            14; Mismatches
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            Pred. No. 15;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   SIMILARITY: BELONGS TO THE RFX FAMILY.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     5.5%; Score 84.5; 33.0%; Pred. No. 15
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            STRAIN=ATCC 824 / DSM 792 / VKM B-1787;
MEDLINE=95050216; PubMed=7961408;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        POTENTIAL
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          PRT;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              01-FEB-1994 (Rel. 28, Created)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     92 168 P
616 AA; 65322 MW;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          EMBL; X85786; CAA59771.1; -.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            Conservative
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          STANDARD;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 RPOD OR SIGA OR CAC1300.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        SCID; Phosphorylation
DNA BIND 92 168
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          T01672; -.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        Query Match
Best Local Similarity
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            SEQUENCE FROM N.A.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            NCB1_TaxID=1488;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          601863; -.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                AND NF-Y.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 209920;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              01-FEB-1994
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 20-AUG-2001
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            33;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          RPSD CLOAB
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 DNA_BIND
SEQUENCE
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 P33656;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     FACTOR
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       RPSD_CLOAB
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                Matches
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              TID DATE OF THE 
                 ŏ
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                qq
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   òλ
```

```
the Buropean Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See http://www.isb-sib.ch/announce/or send an email to license@lsb-sib.ch).
                                                                                                                                                                                                                                                                                                                                                                                                                                                            This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation
                                                                                                                                                                   Noelling J., Breton G., Comelchenko M.V., Makarova K.S., Zeng Q., Gibson R., Lee H.M., Dubois J., Qiu D., Hitti J., Wolf Y.I., Tatusov R.L., Sabathe F., Doucette-Stamm L., Soucaille P., Daly M.J., Bennett G.N., Koonin E.V., Smith D.R.; Forder Bennett G.N., Koonin E.V., Smith D.R.; Genome sequence and comparative analysis of the solvent-producing bacterium Clostridium acetobutylicum."; J. Bacteriol. 183:4813-4818 (2001).

-!- FUNCTION: THE SIGMA FACTOR IS AN INITIATION FACTOR THAT PROMOTES ATTACHMENT OF THE NRA POLYMERASE TO SPECIFIC INITIATION SITES AND PROMOTES THEN IS RELEASED. THIS IS THE PRIMARY SIGMA-FACTOR OF THIS
Sauer U., Treuner A., Buchholz M., Santangelo J.D., Durre P.; "Sportlation and primary sigma factor homologous genes in Clostridium
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   194 ICEATGTPLVFDAHHHVVHDKLPDQEDPS-----VREWVLRARATWQPPEWQVVHL 244
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              109 VRMYLKEIGKVPLLSPEEEIDLAQRIKNGDRSARKKLAEANLRLVVSIAKRYVGRGMLFL 168
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             99 D----AGVRLLMHPEQFIVLNSDRPE-----VRESSVRAMSAHARVMDGLGLARTPWNL 148
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  169 DLIQEGNLGLIKAVEKFDFKKGFKFSTYATWWIRQAITRAIADQARTI-----RIPVHM 222
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 ------LLLHGGKGGRGAELAALIPDLP-DPVRLRLGLENDERAYSPAEL-LP 193
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             223 VETINKLIRVSRQLLQELGREPQPEEIAKIM-DMPVDKVR-----EIMKIAQEPVSLETP 276
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               74; Gaps
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      PRINTS; PR00046; SICHA70_1; 1.
PROSITE; PS00715; SICHA70_1; 1.
PROSITE; PS00716; SIGMA70_2; 1.
Transcription regulation; Sigma factor; DNA-directed RNA polymerase;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      58 IRLYRLSSSLFPML-----DLA----GD-----DTGAAVLTHLAPQLLEAGHAFT 98
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      169 182 POLYMERASE CORE BINDING (POTENTIAL).
339 358 H-T-H MOTIF (BY SIMILARITY).
378 AA; 43260 MW; 13821F827E5D50B8 CRC64;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               Ouery Match 5.5%; Score 84; DB 1; Length 378; Best Local Similarity 22.5%; Pred. No. 8.8;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              95; Indels
                                                                                                                                                                                                                                                                                                                                                                                                                       SIMILARITY: BELONGS TO THE SIGMA-70 FACTOR FAMILY.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         245 SNGIEGPQDRRHSHLIADFPSAYADVPQIEVEAKGK 280
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    328 RFGLDDGRARTLEEVGKEFNVTRERIRQIEAKALRK 363
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              45; Mismatches
                                                                                                                              STRAIN-ATCC 824 / DSM 792 / VKM B-1787;
MEDLINE-21359325; PubMed-11466286;
                                                            Bacteriol. 176:6572-6582(1994).
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               EMBL, AE007642, AAK79271.1; -. PIR, S34307, S34307
HSSP: P00579; 1STG
InterPro: IPR000943; Sigma_70.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            EMBL; Z23080; CAA80625.1; -.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              Conservative
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   Pfam; PF00140; sigma70;
                                                                                                         SEQUENCE FROM N.A.
                                       acetobutylicum.
                                                                                                                                                                                                                                                                                                                                                                                                    BACTERIA
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               DNA-binding.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            62;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           DNA_BIND
SEQUENCE
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      DOMAIN
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   RESULT 13
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 149
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              Matches
g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  a
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               ď
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          δ
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 δ
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   δý
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         δλ
```

271 AA.

PRT;

STANDARD;

TRPA_MYCIT ID TRPA_MYCIT

```
This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation-the Buropean Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See http://www.isb-sib.ch/announce/
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    entities requires a license agreement (See http://www.isb-sib.ch/announce/or send an email to license@isb-sib.ch).
                                                                                                                                                                                                                                                                                                                      Submitted (APR-1998) to the EMBL/JeenBank/DDBD databases.
-!- FUNCTION: THE ALPHA SUBUNIT IS RESPONSIBLE FOR THE ALDOL CLEAVAGE OF INDOLEGLYCEROL PHOSPHATE TO INDOLE AND GLYCERALDEHYDE 3-
                                                                                                                                                                                                                                                                                                                                                                                                      !- CATALYTIC ACTIVITY: L-SERINE + 1-(INDOL-3-YL)GLYCEROL 3-PHOSPHATE
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  ------QLLEAGHAFTDAGVR 103
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    67 ------DGPTIVKATEAALHGGVRVRDTLAAVEAISAAGGHAVVMTYWNPVLR 113
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               104 LLMHPEQFIVLNSDRPEVRESSVRAMSAHARVMD-------GLGLARTPWNLLLL 151
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              Gaps
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          Actinomycetales; Corynebacterineae; Mycobacteriaceae; Mycobacterium.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           Homo sapiens (Human).
Eukaryota; Metazoa: Chordata; Craniata; Vertebrata; Euteleostomi;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            52;
                                                                                                                                                                                                                                                             Alavi M.R., Rouse D.A., Morris S.L., "Nucleotide sequence and functional analysis of the tryptophan synthase genes of Mycobacterium intracellulare.";
                                                                                                                                                                                                                                                                                                                                                                                                                       = L-TRYPTOPHAN + GLYCERALDEHYDE 3-PHOSPRATE + H(2)0.

!- PATHWAY: LAST (FIFTH) STEP IN BIOSYNTHESIS OF TRYPTOPHAN.

!- SUBUNIT: TETRAMER OF TWO ALPHA AND TWO BETA CHAINS (BY
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 5.4%; Score 82.5; DB 1; Length 271; 21.9%; Pred. No. 7.7;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         40; Indels
                                                                                                                                       Bacteria; Firmicutes; Actinobacteria; Actinobacteridae;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       BA0E0C5BC0788318 CRC64;
               30-MAY-2000 (Rel. 39, Created)
30-MAY-2000 (Rel. 39, Last sequence update)
30-MG-2001 (Rel. 40, Last annotation update)
TRYPTOPHAN SYNTHASE ALPHA CHAIN (EC 4.2.1.20).
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 P51693: 000113; Created) 101-0CT-1996 (Rel. 34, Created) 101-NU-1997 (Rel. 38, Last sequence update) 15-JUL-1999 (Rel. 38, Last annotation update) AMYLOID-LIKE PROTEIN 1 PRECURSOR (APLP).
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           -! - SIMILARITY: BELONGS TO THE TRPA FAMILY.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         19; Mismatches
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                64 SSSLFPMLDLAGDDTGAAVLTHLAP------
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               InterPro; IPR003009; FMN_enzyme.
InterPro; IPR002028; TRP_synthase_alpha.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   Pfam; PF00290; trp_syntA; 1.
ProDom; PD001535; TRP_synthase_alpha; 1.
PROSITE; PS00167; TRP_SYNTHASE_ALPHA; 1.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          152 HG-----GKGGRGAELAALIPD 168
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 114 YGIDAFARDLASAGGYGLITPDLIPD 139
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      PRT;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               Tryptophan blosynthesis; Lyase.
SEOUENCE 271 AA; 27804 MW;
                                                                                                                   Mycobacterium intracellulare.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              EMBL; AF057042; AAC17135.1; -
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  21.9%;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         32; Conservative
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      STANDARD;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  Best Local Similarity
Matches 32; Conserva
                                                                                                                                                                                                                        SEQUENCE FROM N.A.
                                                                                                                                                                               NCBI_TaxID=1767;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       SIMILARITY).
                                                                                                                                                                                                                                                                                                                                                                                       PHOSPHATE.
                                                                                                                                                                                                                                           STRAIN-BATTY;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      APP1_HUMAN
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     SEQUENCE
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               Query Match
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          APLP1.
                                                                                                    PRPA.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              HUMAN
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   APP1
ID
DE
DE
GN
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             ΟŻ
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    qq
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          Qγ
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               q
```

us-09-724-296-39.rsp

```
δy
                                                                                                       D
                                                                                                                                                                  Ω
                                                                                                                                                                                                                       g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See http://www.isb-sib.ch/announce/
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                12;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    Hum. Genet. 102:192-196(1998).
-!- SUBCELLULAR LOCATION: TYPE I MEMBRANE PROTEIN. C-TERMINALLY
PROCESSED IN THE GOLGI COMPLEX AND IS THEN SECRETED.
-!- TISSUE SPECIFICITY: PREDOMINANTLY EXPRESSED IN BRAIN, PARTICULARLY
IN THE CEREBRAL CORTEX POSTSYNAPTIC DENSITY.
                                                                                                                      Paliga K., Peraus G., Kreger S., Duwrrwang U., Hesse L., Multhaup G., Masters C.L., Beyreuther K., Weidemann A.; Human amyloid precursor-like protein 1-cDNA cloning, ectopic expression in COS-7 cells and identification of soluble forms in the cerebrosphan fluid.";
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            148 LLLLHGGKGGRGAE-LAALIPDLPDPVRLRLGLENDERAYSPAELLPICEATGTPLVFDA 206
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 84 CLRDPQRVLEYCRQMYPELQIARVEQATQAIPMERWCGGSRSGSCAHPHHQVVPFRCLPG 143
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         95 HAFTDAGVRLLMHPEQFIVLNSDRPEVRESSVR-----AMSAHARVMDGLGLARTPWN 147
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      Gaps
                                                                                                                                                                                                                                                                                                                                                                                                                          Olsen A., Tryggvason K.; "Structure of the human amyloid-precursor-like protein gene \ensuremath{\mathsf{APLPl}} at
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        53 CAAHDIRLYRLSSSLFPMLDLA-------GDDTGAAVI.TH--LAPQLLEAG 94
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     N-LINKED (GLCNAC. ..) (POTENTIAL).
N-LINKED (GLCNAC. ..) (POTENTIAL).
N-LINKED (GLCNAC. ..) (POTENTIAL).
                                                                                                                                                                                                                                                                                                                                                                                             Lenkkeri U., Kestila M., Lamerdin J., McCready P., Adamson A.,
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           DB 1; Length 650;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  CLATHRIN-BINDING (POTENTIAL).
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  AMYLOID-LIKE PROTEIN 1.
EXTRACELLULAR (POTENTIAL).
POTENTIAL.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        -> A (IN REF. 2).
121A034B708C67CA CRC64;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        CYTOPLASMIC (POTENTIAL)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                25; Mismatches
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           5.4%; Score 82.5; 1
20.2%; Pred. No. 23;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  -!- PTM: N- AND O-GLYCOSYLATED.
-!- SIMILARITY: BELONGS TO THE APP FAMILY.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                or send an email to license@isb-sib,ch).
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           POTENTIAL
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             POLY-GLU.
POLY-GLU.
                                                                                                                                                                                                                                                                               Eur. J. Biochem. 250:354-363(1997).
                                                                                                       MEDLINE=98088960; PubMed=9428684;
                                                                                                                                                                                                                                                                                                                                                                     MEDLINE=98180887; PubMed=9521588;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         <u>=</u>
-
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         PERUTY, AGENTA, AGENTRA; DERINTS; PRO177; AGENTRA; DEN CONTRA; SMARY; SMOROS; AGENTRA; DER COSTE; PSO0319; AGENTRA; DER COSTÉ; PSO0320; PAGENTRA; DER COSTÉ; PSO0320; PS
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                EMBL; AD000864; AAB50173.1; -
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           48 P
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                Glycoprotein; Transmembrane;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     EMBL; U48437; AAB96331.1; -.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     InterPro; IPR001868; A4_APP
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   Conservative
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     650
580
603
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        268
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  650 AA;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         Similarity
                                                                                  SEQUENCE FROM N.A.
                                                                                                                                                                                                                                                                                                                                       SEQUENCE FROM N.A.
                          NCBI_TaxID=9606;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     19q13.1.";
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                DOMAIN
TRANSMEM
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              Query Match
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           CARBOHYD
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                CARBOHYD
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        SEQUENCE
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      CARBOHYD
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           CONFLICT
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            Local
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                DOMAIN
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           SIGNAL
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           DOMAIN
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              DOMAIN
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           CHAIN
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                Matches
δ
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 q
```

```
SEQUENCE FROM N.A.

STRAIN=K12 / W3110;

Takemoto K., Mori H., Murayama N., Kataoka K., Yano M., Itoh T.,

Yamamoto Y., Inokuchi H., Miki T., Hatada E., Fukuda R., Ichihara S.,

Mizuno T., Makino K., Nakata A., Yura T., Sampei G., Mizobuchi K.;

"Systematic sequencing of the Escherichia coli genome: analysis of the
"Systematic sequencing of the Escherichia coli genome: analysis of the
4.0 - 6.0 min (189,987 - 281,416bp) region.";

Submitted (FEB-1996) to the EMBL/GenBank/DDBJ databases.
                                      207 HHHVVHDKLPDQEDPSVREWVLRARATWQPPEWQVVHLSNGIEGPQDRRHSHLIADFPSA 266
                                                                          -----SPSTRSW----SPP----GSRVEGAEDEEEEE---SFPQP 252
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         Blattner F.R., Plunkett G. III, Bloch C.A., Perna N.T., Burland V., Riley M., Collado-Vides J., Glasner J.D., Rode C.K., Mayhew G.F., Gregor J., Davis N.W., Kirkpatrick H.A., Goeden M.A., Rose D.J., Mau B., Shao Y.;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         Schramm S., Duncan M., Allen E., Araujo R., Aparicio A., Chung E., Davis K., Federspiel N., Hyman R., Kalman S., Komp C., Kurdi O., Lashkari D., Lew H., Lin D., Namath A., Oefner P., Roberts D., Davis R.W.;
                                                                                                                                                                                                                                                                                                                                                                                                             Bacteria; Proteobacteria; gamma subdivision; Enterobacteriaceae;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       Li S.J., Cronan J.E. Jr., "Growth rate regulation of Escherichia coli acetyl coenzyme A carboxylase, which catalyzes the first committed step of lipid
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       "The complete genome sequence of Escherichia coli K-12."; Science 277:1453-1474(1997).
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    Submitted (DEC-1995) to the EMBL/GenBank/DDBJ databases.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       MEDLINE-88058791; PubMed-3316192;
Tomasiewicz H.G., McHenry C.S.;
"Sequence analysis of the Escherichia coli dnaE gene.";
J. Bacteriol. 169:5735-5744(1987).
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              Submitted (SEP-1996) to the EMBL/GenBank/DDBJ databases
                                                                                                                                                                                                                                                                                                            01-MAR-1989 (Rel. 10, Last sequence update)
20-AUG-2001 (Rel. 40, Last annotation update)
DNA POLYMERASE III ALPHA SUBUNIT (EC 2.7.7.7)
                                                                                                                                                                                                                                                      PRT; 1160 AA
193 MLLPCGSDRFRGVEYVCCPPPGTPDPSGTAVG-
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      SEQUENCE FROM N.A.
STRAIN-K12 / MC1655;
MEDLINE-97426617; Pubmed-9278503;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                SEQUENCE OF 1070-1160 FROM N.A. MEDLINE=93123150; Pubmed=7678242;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              MEDLINE=92246902; PubMed=1575709;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     Bacteriol. 175:332-340(1993).
                                                                                                                  267 ----YADVPQIEVEAKGKEEAI 284
                                                                                                                                                     253 VDDYFVEPPQAEEE----EETV 270
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         SEQUENCE OF 1137-1160 FROM N.A.
                                                                                                                                                                                                                                                                                          01-MAR-1989 (Rel. 10, Created)
                                                                                                                                                                                                                                                      STANDARD;
                                                                                                                                                                                                                                                                                                                                                                    DNAE OR POLC OR B0184.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              STRAIN=K12 / W3110;
Yamamoto Y.;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               SEQUENCE FROM N.A.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       SEQUENCE FROM N.A.
                                                                                                                                                                                                                                                                                                                                                                                             Escherichia coli
                                                                                                                                                                                                                                                                                                                                                                                                                                                 NCBI_TaxID=562;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 O'Donnell M.;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 biosynthesis.
                                                                                                                                                                                                                                                                                                                                                                                                                             Escherichia
                                                                                                                                                                                                                                                      DP3A_ECOLI
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            REVIEW.
                                                                                                                                                                                                                                                                        P10443
                                                                                                                                                                                                                                   DP3A_ECOLI
```

```
This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation - the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See http://www.isb-sib.ch/announce/or send an email to license@isb-sib.ch).
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               N PYROPHOSPHATE + DNA (N.)
SUBUNIT: CONTAINS A CORE (COMPOSED OF ALPHA, EPSILLON, AND THETA CHAINS) THAT ASSOCIATES WITH A TAU SUBUNIT WHICH ALLOW THE CORE DIMERIZATION TO FORM THE POLIII' COMPLEX. POLIII' ASSOCIATES WITH FIRE GAMMA COMPLEX (COMPOSED OF CHAINS GAMMA, DELTA, DELTA', PSI, AND CHI) AND WITH THE BETA CHAIN. THE FINAL COMPOSITION OF THE COMPLEX IS: (ALPHA, EPSILON, THETA) [2]-TAU[2]-(GAMMA, DELTA', DELTA', PSI, CHI) [2]-BETA[4].
SUBCELLULAR LOCATION: CYTOPLASMIC.
"Accessory protein function in the DNA polymerase III holoenzyme from
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     158 RGAEL----AALIPD------LPDPVRL---RLGLENDERAYSPAELLPICEATGTPL 202
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             99 DAGVRLLMHPEQFIVLNSDRPEVRESSVRAMSAHARVMDGLGLARTPWNLLLLHGGK-GG 157
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        110 -------GAAGPIIDRDWLIELNEGLILLEGGRMGD 138
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                139 VGRSLLRGNSALVDECVAFYEEHFPDRYFLELIRTG-RPDEESYLHA-AVELAEARGLPV 196
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        203 V------FDAH--HHVVHD--KLPDQEDPSVREWVLRARATWQPPEWQVVHLSN 246
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                Gaps
                                                                                                                                                       polymerase III: identification of the responsible mutations and alignment with other DNA polymerases."; Genetics 134:1039-1044(1993).
-!- FUNCTION: DNA POLYMERASE III IS A COMPLEX, MULTICHAIN ENZYME RESPONSIBLE FOR MOST OF THE REPLICATIVE SYNTHESIS IN BACTERIA. THIS DNA POLYMERASE ALSO EXHIBITS 3' TO 5' EXONUCLEASE ACTIVITY.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              39 YSSNIKTLRGAADYCAAHDIRLYRLSSSLFPMLDLAGDDTGAAVLTHLAPQLLEAGHAFT 98
                                                                                             MEDIINE-93387658; PubMed=8375647;
Fijalkowska I.J., Schaaper R.M.;
"Antimutator mutations in the alpha subunit of Escherichia coli DNA
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            93;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 5.4%; Score 82.5; DB 1; Length 1160; 22.5%; Pred. No. 48;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          SMART; SM00481; POLIIIAc; 1.
Transferase; DNA-directed DNA polymerase; DNA replication;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            84; Indels
                                                                                                                                                                                                                                                                                THE ALPHA CHAIN IS THE DNA POLYMERASE.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      1160 AA; 129904 MW; 1A4F75F373841716 CRC64;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          30; Mismatches
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   U70214; AAB08613.1; -. S52931; AAB24889.1; -. M96394; AAA70369.1; ALT_SEQ. D49445; BAA08424.1; -.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           InterPro; IPR003141; PHP_N.
InterPro; IPR002309; tRNA-synt_2.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          EMBL; M19334; AAC36920.1; -.
EMBL; AE000127; AAC73295.1; -.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 InterPro; 1PRUVALON, 1.
Pfam; PF01231; PHP_N; 1.
Pfam; PF01336; tRNA_anti; 1.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     Best_Local Similarity 22.5%;
Matches 60; Conservative 3
                                      Bioessays 14:105-111(1992).
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      D83536; BAA77859.1;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      EcoGene; EG10238; dnaE
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                PIR; C28390; DJEC3A.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         PIR; A40637; A40637
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     Complete proteome SEQUENCE 1160 A
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          SUBFAMILY.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    Query Match
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           EMBL;
  Op
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         ÓΥ
```

```
Db 197 VATNDVRFIDSSDFDAHEIRVAIHDGFTLDDPKRP------RNYSPQQYM----- 240
Qy 247 GIECPQDRRHSHLIADFPSAYADVPQI 273
Db 241 ----RSEEEMCELFADIPEALANTVEI 263
Search completed: January 15, 2002, 14:07:28
JOB Lime: 957 sec
```

```
; Search time 57.72 Seconds
(without alignments)
114.522 Million cell updates/sec
                                                                                                                                                                                                                                                                               1 QLGLVCLTVGPEVRFRTVTL.....VEAKGKEEAIAALRLMAPFK
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          Issued_Patents_AA:*
!: /cgn2_6/ptodata/2/iaa/5A_COMB.pep:*
!: /cgn2_6/ptodata/2/iaa/5B_COMB.pep:*
!: /cgn2_6/ptodata/2/iaa/6A_COMB.pep:*
!: /cgn2_6/ptodata/2/iaa/6B_COMB.pep:*
!: /cgn2_6/ptodata/2/iaa/PCTUS_COMB.pep:*
!: /cgn2_6/ptodata/2/iaa/PcTUS_COMB.pep:*
!: /cgn2_6/ptodata/2/iaa/PcTUS_COMB.pep:*
                           Compugen Ltd.
                                                                                                                                                                                                                                                                                                                                                                                                                                                  Total number of hits satisfying chosen parameters:
GenCore version 4.5
Copyright (c) 1993 - 2000 Comp
                                                                                                                                                                                                                                                                                                                                                                                                         212252 seqs, 22503292 residues
                                                                                                                                         January 15, 2002, 13:51:25
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries
                                                                                           - protein search, using sw model
                                                                                                                                                                                                                                                                                                                                     BLOSUM62
Gapop 10.0 , Gapext 0.5
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     Minimum DB seq length: 0
Maximum DB seq length: 200000000
                                                                                                                                                                                                                                       US-09-724-296-39
                                                                                                                                                                                                                                                           Perfect score:
                                                                                                                                                                                                                                                                                                                                     Scoring table:
                                                                                           OM protein
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            Database :
                                                                                                                                                                                                                                                                                    Sequence:
                                                                                                                                                                                                                                                                                                                                                                                                            Searched:
                                                                                                                                         Run on:
```

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

Sequence 4, Appli Sequence 6, Appli Sequence 6, Appli Sequence 5, Appli Sequence 2, Appli Patent No. 5290690 Sequence 107, App Sequence 2, Appli Sequence 2, Appli Sequence 17, Appli Sequence 17, Appl Sequence 16, Appl Sequence 3, Appli Sequence 3, Appli Sequence 27, Appl Appl Appl Sequence 15, Appl. Sequence 15, Appl. Appli Appli Appli Appli Appli Appli Appli Sequence 12, P Sequence 6, Ap Sequence 4, Ap Sequence 3, 1 Sequence 1, 1 Sequence 1, 1 Sequence 5, 1 Sequence 9, Sequence 5, Description Sequence US-08-633-879C-2 US-08-339-152A-17 US-08-139-152A-16 US-08-007-999B-3 US-08-861-774E-27 US-08-861-774E-27 US-08-861-774E-27 US-08-804-138-6 US-08-439-009A-4 US-09-036-987A-6 US-09-036-987A-6 US-08-472-534-5 US-08-472-534-5 US-08-375-709-15 US-08-752-929-15 US-09-090-793-9 US-08-804-227C-6 US-09-413-814-107 PCT-US93-03035-2 US-09-335-409-5 US-07-637-399-1 US-07-640-476-5 -09-428-517-3 SUMMARIES 5290690-2 Query Match Length DB Score 85.5 85.5 85.5 84.5 83 82.5 82.5 81.5 81.5 81 80.5 80 80 79.5 79.5 79.5 79 78.5 78.5 78.5 Result . №

Sequence 1, Appli Sequence 2, Appli Sequence 2, Appli Sequence 4, Appli Sequence 59, Appl Sequence 59, Appl Sequence 59, Appl Sequence 6, Appli Sequence 7, Appli Sequence 2, Appli Sequence 2, Appli Sequence 3, Appli Sequence 3, Appli Sequence 3, Appli Sequence 42, Appli Sequence 42, Appli Sequence 42, Appli Sequence 42, Appli Sequence 2, Appli Sequence 2, Appli	Sequence 2, Appli
US-08-112-703-1 US-07-642-734C-2 US-08-439-003A-2 US-09-428-517-4 US-08-931-608A-4 US-08-95-028-59 US-09-95-028-59 US-09-95-028-59 US-09-413-814-67 US-09-413-814-67 US-08-804-227C-2 US-08-804-227C-3 US-08-404-517A-97 US-08-804-227C-3 US-08-804-227C-3 US-08-804-227C-3 US-08-804-227C-3 US-08-804-227C-3 US-08-804-227C-3	US-09-051-019-2
11 0 6 4 4 4 4 4 4 6 0 4 0 0 4 0 0 0	ĸ
39491 3491 3519 3519 3519 719 719 719 719 719 719 719 719 719 7	5289
	4. 9.
78.5 8.5 8.5 77 77 77 77 77 77 76 76 76	6.57
20 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0	4 C

ALIGNMENTS

```
APPLICANT: ABBOTT LABORATORIES
TITLE OF INVENTION: PURIFIED THERMOSTABLE ENDONUCLEASE
                                                                                                                                                                                                                                                                                                                                                             PatentIn Release #1 0, Version #1.25
                                                                                                                                                            ADDRESSEE: ABBOTT LABORATORIES
STREET: D-377 AP6D, ONE ABBOTT PARK ROAD
CITY: ABBOTT PARK
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    REGISTRATION NUMBER: 32,459
REFERENCE/DOCKET NUMBER: 5145.PC.01
TELECOMMUNICATION INFORMATION:
TELEPHONE: 708-937-4884
                                                                                                                                                                                                                                                                                                                                              OFENAL:
SOFTWARE: Patentin.
CURRENT APPLICATION DATA:
APPLICATION NUMBER: PCT/US93/03035
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      ALTORATION NUMBER: US/07/860,702 FILING DATE: 31 MAR-1992 ATTORNEY/ACTORIT INFORMATION:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                APPLICATION NUMBER: US/07/869,306
FILING DATE: 16-APR-1992
                                                                                                                                                                                                                                                                                                                   IBM PC compatible
SYSTEM: PC-DOS/MS-DOS
                                       Sequence 2, Application PC/TUS9303035 GENERAL INFORMATION:
                                                                                                                                                                                                                                                                                                      Floppy disk
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       BRAINARD, THOMAS D
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       270 amino acids
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  SEQUENCE CHARACTERISTICS:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  TELEPHONE: 708-937-488
TELEFAX: 708-937-2623
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           INFORMATION FOR SEQ ID NO:
                                                                                                                                                                                                                                                                                                                                                                                                                     FILING DATE: 19930330
CLASSIFICATION:
PRIOR APPLICATION DATA:
                                                                                                                                                                                                                                                                               COMPUTER READABLE FORM:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   FILING DATE: 16-APR-1
PRIOR APPLICATION DATA:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           ; MOLECULE TYPE: protein PCT-US93-03035-2
                                                                                                                     NUMBER OF SEQUENCES: 2
CORRESPONDENCE ADDRESS:
                                                                                                                                                                                                                                                                                                                                          OPERATING SYSTEM:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        AMINO ACID
                                                                                                                                                                                                                                        COUNTRY: USA
ZIP: 60064-3500
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               linear
                                                                                                                                                                                                                       ILLINOIS
                                                                                                                                                                                                                                                                                                      MEDIUM TYPE:
  RESULT 1
PCT-US93-03035-2
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          TOPOLOGY:
                                                                                                                                                                                                                                                                                                                       COMPUTER:
                                                                                                                                                                                                                       STATE:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       NAME:
RESULT
```

17; Gaps 89; 6.2%; Score 94.5; DB 5; Length 270; 27.5%; Pred. No. 0.017; tive 29; Mismatches 67; Indels 89 Query Match
Best Local Similarity 27.55
Matches 70; Conservative

276 EAK 278

452 PAE 454

Qy Db

```
280 NHDGASSGLTVPNGSSQEIVLKRALADAGCAASSVGYVEAHGTGTTLGDPIEIQALNAVY 339
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            398 GDLRLTVTRARTPW--PDWNTPRRACVSSFGMSG----TNAHVVLEEAPAATCTPPAPER 451
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          178 GLENDERAYSPAELLPICEATGTP-----LVFDAHHHVVHDKLPDQE-DPSV-- 223
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    340 GLGRD--VATPLLIGSVKTNLGHPEYASGITGLLKVVLSLQHGQIPAHLHAQALNPRISW 397
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           224 ---REWVLRARATWQPPEWQVVH----LSNGIEGPQDRRHSHLI-ADFPSAYADVPQIEV 275
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          Query Match

5.9%; Score 90; DB 3; Length 1421;
Best Local Similarity 21.5%; Pred. No. 0.78;
Matches 78; Conservative 37; Mismatches 122; Indels 126; Gaps
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 180 AVDT---AYSSSL-----VAVHLACQSLRSGECSTALAGGVSLMLSPSTLV 222
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           113 VLNSDRPEVRESSVRAMSAHARVMDGLGLARTPWNLLL-·LHGGK--GGR------ 158
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  223 WLSKTRALARDGRCKAFSAEA---DGFGRGEGCAVVVLKRLSGARADGDRILAVIRGSAI 279
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            ----LPDPVRLR----L 177
                                                                                                                                                                                                                                                                                                                                                                                    167 ---CLDTCHA-----YAAGYDVAEDPL-----GVLDALDRAVGLERVP---VVHLNDS 208
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   129 VFIGIGP-----SEYEAALPQATASAEIDAHGGLGTMPSVGAGRISYALGLRGPCV 179
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             55 AHDIRLYRLSSSLFPMLDLAGDDTGAAVLTHLAPQLLEAGHAFT--DAGVRLLMHPEQFI 112
                                                                                                                                                                                                                                                                         188 PAELLPICEATGTPLVFDAHHHVVHDKLPDQEDPSVREWVLRARATWQPPEWQVVHLSNG 247
                                                                                                              76 DDTGAAVLTHLAPQLLEAGHAFTDAGVR-LLMHPEQFIVLNSDRPE-VRESSVRAMSAHA 133
                                                                                                                                                                 ------GVEYVVHP-----GSGRPERVKEGALKALRL-- 127
                                                                                                                                                                                                                                  134 RVMDGLGLARTPWNLLLLH----GGK--GGRGAELAALIPDLPDPVRLRLGLENDERAYS 187
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            5 VCLTVGPEVRFRTVTLSRYRALSPAEREAKLLDLY------SSNIKTLRGAADYCA 54
22 RYRALSPAEREAKLLDLYSSNIKTLRGAADYCAA--HDIRLYRLSS--SLF--PMLDLAG 75
                                                   42 RPRALSPAEVEA-----FRALREASGGLPAVIHASYLVNLGAEGELWEKSVASLAD 92
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          APPLICANT: Schupp, Thomas
APPLICANT: Ligon, James
APPLICANT: Molnar, Istvan
APPLICANT: Zirkle, Ross
APPLICANT: Zirkle, Ross
APPLICANT: Cyr, Devon
APPLICANT: Goellach, Joern
TITLE OF INVENTION: GENES FOR THE BIOSYNTHESIS OF EPOTHILONES
FILE REFERENCE: 4-30582A
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      CURRENT APPLICATION NUMBER: US/09/335,409 CURRENT FILING DATE: 1999-06-17 NUMBER OF SEQ ID NOS: 30 SOFTWARE: PatentIn Ver. 2.0
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              Sequence 2, Application US/09335409 Patent No. 6121029 GENERAL INFORMATION:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                ; ORGANISM: Sorangium cellulosum US-09-335-409-2
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  159 ---GAELAALIPD-----
                                                                                                                                                                                                                                                                                                                                                                                                                                                                       248 IEGPQDR--RHSHLI 260
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              209 VGGLGSRVDHHAHLL 223
                                                                                                                                                                              93 DLEKAALL----
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            1421
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               US-09-335-409-2
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                SEQ ID NO 2
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          TYPE: PRT
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         QQ
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   Óγ
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            QQ
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      δy
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    QY
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           q
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        ò
                                                                                                                                                                              qq
                                                                                                                                                                                                                                                                                                                                                         ò
                                                                                                                                                                                                                                                                                                                                                                                                                qq
                                                                                                                                                                                                                                                                                                                                                                                                                                                                          òχ
```

```
54 EAKLAKIKSWASKMEALTSKSAADPEGY-LAHPVNAYKLVKRLNTDWPALGDLVLQDASA 112
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     81 AVLTHLAPQLL-----EAGHAFTDAGVRLLMHPEQFIVLNSD---RPEVRESSVRAMS 130
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                113 GFVANESVQRQFFFPTDEDESGAA-----RALMRLQDTYKLDPDTISRGELPGTKYQAML 166
                                                                                                                              APPLICANT: Helaakoski, Tarja I.
APPLICANT: Annuen, Pia P.
APPLICANT: Nisvi, Ritva W.
APPLICANT: No. 5928922elainen, Minna K.
APPLICANT: No. 5928922elainen, Minna K.
TITLE OF INVENTION: 2 SUBUNIT OF PROLYL-4-HYDROXYLASE
TITLE OF INVENTION: NUCLEIC ACID SEQUENCES ENCODING SUCH SUBUNIT AND
TITLE OF INVENTION: METHODS FOR PRODUCING THE SAME
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         131 AHARVMDGLGLARTPWN-------157
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                5.8%; Score 88; DB 2; Length 537;
22.2%; Pred. No. 0.28;
tive 39; Mismatches 116; Indels 160; Gaps
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             32 EAKLLDLYSSNIK----TLRGAAD---YCAAHDIRLYRLSSSL---FPML-DLAGDDTGA 80
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          OPERATING SYSTEM: Windows SOFTWARE: FastSEQ for Windows Version 2.0b CURRENT APPLICATION DATA:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             8389-0041-999
                                                                                                                                                                                                                                                                                                                                                         STREET: 1155 Avenue of the Americas
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            US/08/633,879C
                                                                                                                                                                                                                                                                                                                                 Pennie & Edmonds, LLP
              Sequence 2, Application US/08633879C Patent No. 5928922
                                                                                 APPLICANT: Kivirikko, Kari I.
APPLICANT: Pihlajaniemi, Taina
APPLICANT: Helaakoski, Tarja I.
APPLICANT: Annunen, Pia P.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             30,605
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             REFERENCE/DOCKET NUMBER: 83: TELECOMMUNICATION:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           INFORMATION FOR SEQ ID NO: 2:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       IBM Compatible
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        ATTORNEY/AGENT INFORMATION:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             1.0-APR-1996
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      NE: 650-493-4935
: 550-493-5556
66141 PENNIE
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           537 amino acids
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            Query Match 5.8%
Best Local Similarity 22.2%
Matches 90; Conservative
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     Abrams, Samuel B
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   SEQUENCE CHARACTERISTICS:
                                                                                                                                                                                                                                                                                              50
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        internal
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               Diskette
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      single
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                CLEANS CANAL STATE CLASSIFICATION: APPLICATION DATA: APPLICATION NUMBER:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 MOLECULE TYPE: protein FRAGMENT TYPE: internal
                                                                                                                                                                                                                                                                                                                                                                                                                                               ZIP: 10036-2811
COMPUTER READABLE FORM:
MEDIUM TYPE: Diskett
                                                                                                                                                                                                                                                                                                                 CORRESPONDENCE ADDRESS:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          REGISTRATION NUMBER:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               APPLICATION NUMBER:
                                                                                                                                                                                                                                                                                           NUMBER OF SEQUENCES:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                amino acid
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              linear
                                                                GENERAL INFORMATION:
                                                                                                                                                                                                                                                                                                                                                                                 New York
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      STRANDEDNESS:
                                                                                                                                                                                                                                                                                                                                                                                                                              USA
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  FILING DATE:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            FILING DATE:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                TELEPHONE:
                                                                                                                                                                                                                                                                                                                                           ADDRESSEE:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              COPOLOGY:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            US-08-633-879C-2
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         COMPUTER:
US-08-633-879C-2
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       TELEFAX:
                                                                                                                                                                                                                                                                                                                                                                                                                                 COUNTRY:
                                             . Patent No.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    qq
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               ó
```

Db 172 MLLPCGSDRFRGVEXVCCPP 203	208 HHVVHDKLPDQEDPSVREWVLRARATWQPPEWQVVHLSNGIEGPQDRRHSHLIADFPSA- 	Db 204DPSTRSWPLGGRA	233 DDYFVEPN, ''EEEEEEE	RESULT 5	US-08-339-152A-16 ; Sequence 16, Application US/08339152A ; Patent No. 5643726	Rudolph E. Dora M.	; TITLE OF INVENTION: Methods For Modulating Transcription ; TITLE OF INVENTION: From The Amyloid -Protein Precursor (APP) Promot ; NUMBER OF SEQUENCES: 33	CORRESPONDENCE ADDRESS: ADDRESSEE: STREET: 1100 New York Ave., NW, Suite 600	: Washi E: DC TRY: US	PE PE	COMPUTER: 1BM PC compatible COMPUTER: 1COMPATING SYSTEM: PC-DOS/MS-DOS COFTWARE: PatentIn Release #1.0, Version #1.25	; CURRENT APPLICATION DATA: ; APPLICATION NUMBER: US/08/339,152A ; FILING DATE: 10-NOV-1994	CLASSIFICATION: 435 : ATTONERY/AGENT INFORMATION: : NAWE: Steffe, Eric K.	; REGISTRATION NUMBER: 36,688 ; REPERENCE/SOCKET NUMBER: 0609.4120000 ; THE ECOMMINICATION INDICATION	;	; TELEX: ; INFORMATION FOR SEQ ID NO: 16; ; SEQUENCE CHARACTERISTICS:	amino ac acid inear	; MOLECULE TYPE: protein US-08-339-152A-16	r Match 5.6%; Score 85.5	Sımılatıly 55; Conservat	OY 53 CAAHDIRLYRLSSSLEPMLDLAGDDTGAAVLTHLAPQLLEAG 94	OY 95 HAFTDAGVRLLMHPEQFIVLNSDRPSVRESSVRAMSAHARVMDGLGLARTPWN 147	Db 143 EFVSEALLVPEGCRFLHQERMDQCESSTRRHQEAQEACSSQGLILHGSG 191	QY 148 LLLLHGGGGGGELAALIPDLPDPVRLRLGLENDERAYSPAELLPICEATGTPLVFDAH 207	208 HHVVHDKLPDOEDPSVREWULRARATWOPPEWOVVHLSNGIEGPODRRHSHLIADFPSA	
Oy 158RGAELAALIPDLPDPVRLKLGLENDERAYS	224 DLHRAVELTRRLLSL-DPSHERAGGNLRYFERLLEBERGKSLSNQTDAGLATQENLYERP	Oy 189 AELLPICEATGTPLVEDAHH	212 -HDKLPDQEDPSVREWVLRARATWQPPEWQVVHLSNGIEGPQDR	EIERIKEIAKPKL	Qy 255 RHSHLIADFPSAYADVPQIEVEAKGKEEAIAAL 287 :	RESULT 4	US-UB-339-12ZA-1/ ; Sequence 17, Application US/0833915ZA ; Patent No. 56437Z6 · CENTRAL INFORMATION.	CENERAL INCRALION: APPLICANT: Tanzi, Rudolph E. APPLICANT: Kovacs, Dora M.	; TITLE OF INVENTION: Methods For Modulating Transcription ; TITLE OF INVENTION: From The Amyloid -Protein Precursor (APP) Promoter ; COBBEGGRANDENTS: 33	; CORRESPONDENCE ADDRESS: ; ADDRESSE: STERNE, KESSLER, GOLDSTEIN & FOX P.L.L.C. ; STREET: 1100 New York Ave., NW, Suite 600	E	O 46 A	; COMPUTER: IBM PC compatible ; OPERATING SYSTEM: PC-DOS/MS-DOS ; SOFTWARE: PatentIn Release #1.0, Version #1.25	; CURRENT APPLICATION DATA: ; APPLICATION NUMBER: US/08/339,152A ; FTI,NG DATE: 10-NOV-1994		; NAME: Steffe, Eric K. ; REGISTRATION NUMBER: 36,688 ; REFERENCE/DOCKET NUMBER: 0609.4120000	; TELECOMMUNICATION INFORMATION: ; TELEPHONE: 202-371-2600 ; TELEFAX: 202-371-2540		; LENGTH: 634 amino acids ; TYPE: amino acid · monorogy. I spear	7	atch 5.6%; Score 85.5; DB 1; Length 634:	Matches 55; Conservative 22; Mismatches 83; Indels 99; Ga	OY 53 CAAHDIRLYRLSSSLFPWLDLAGDDTGAAVLTHLAPQLLEAG 94	95 HAFTDAGVRLIMHPEQFIVLNSDRPEVRESSVRAMSAHARVMDGLGLARTPWN	\sim	Qy 148 LLLLHGGKGGRGAELAALIPDLPDPVRLRLGLENDERAYSPAELLPICEATGTPLVFDAH 207

```
10;
                                                                                                                                                                                        dulating Transcription
id -Protein Precursor (APP) Promoter
SWQVVHLSNGIEGPODRRHSHLIADFPSA- 266
| | | : | | : | | | ...
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  SVR------AMSAHARVMDGLGLARTPWN 147
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              EWQVVHLSNGIEGPQDRRHSHLIADFPSA- 266
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             SLENDERAYSPAELLPICEATGTPLVFDAH 207
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    35.5; DB 1; Length 653;
Wo. 0.73;
matches 83; Indels 99; Gaps
                                                                                                                                                                                                                                       STEIN & FOX P.L.L.C.
Suite 600
                                                                                                                                                                                                                                                                                                                                                          Version #1.25
                                                                                                                                                                                                                                                                                                                                                                                                                                                    20000
```

252 OY 208 HHVVHDKLPDQEDPSVREWVLRARATWQPPEWQVVHLSNGIEGPQDRRHSHLIADFPSA-		Qy 267YADVPQIEVEAKGKEE 282	Db 253 DDYFVEPPQABEBEBEBE 271		US-08-276A-3	; Sequence 3, Application US/UBOR92/0A ; Patent No. 5891591	; GENERAL INFORMATION:	 	3 APPLICANT: Solomon, Frank 3 TITLE OF INVENTION: AMYLOID PRECURSOR-LIKE PROFEIN AND USES	(ADDRESSEE: Sterne, Kessler, Goldstein & Fox P.L.L.C.	100 New York Ave., hington	; STATE: DC ; COUNTRY: USA	; ZIP: 20005 ; COMPUTER READABLE FORM:	MEDIUM TYPE: Floppy disk	SYSTEM: PC-DOS/MS-DOS	-, -	; APPLICATION NUMBER: US/08/689,276A ; FILING DATE: 06-AUG-1996	; PRIOR APPLICATION DATA: : APPLICATION NIMPER: IIS 08/007.999	; FILING DATE: 21-JAN-1993	, APPLICATION NUMBER: US 07/872,642	; FILING DATE: 20-APR-1992 ; PRIOR APPLICATION DATA:	; APPLICATION NUMBER: US 07/930,022 ; FILING DATE: 17-AUG-1992	; ATTORNEY/AGENT INFORMATION: ; NAME: JORGE A. GOLDSTEIN	REGISTRATION NUMBER: 29,021 REFERENCE/DOCKET NUMBER: 0609,3520003	Υ	; TELEFAX: (202)371-2540 ; INFORMATION FOR SEQ ID NO: 3:	; SEQUENCE CHARACTERISTICS: ; LENGTH: 653 amino acids	; TYPE: amino acid ; TOPOLOGY: linear	; MOLECULE TYPE: protein ups 10; US-08-689-276A-3	atch 5.6%; Score 85.5; DB 2; Length 653; cal Similarity 21.2%; Pred. No. 0.73;	Matches 55; Conservative 2 Qy 53 CAAHDIRLYRLSSSLFPMLDLA-	
224DPSTRSWPLGGRA	Qy 267xadvpQIEVEAKGKEE 282	Db 253 DDYFVEPPQAEGEGEGEE 271		RESULT 6 05-08-07-099B-3	Sequence of Application os/cook/3336 Patent No. 5851787 CEMPRAL INTOBARATION.	APPLICANT: Wasco, Wilma	APPLICANT: Bupp, Keith APPLICANT: Macendantz. Margaret	TITLE OF INVANTION: AMYLOID PRECURSOR-LIKE PROTEIN AND USES	6	CONTRACTOR AND TABLES SEEN SEASON OF THE FOX, P.L.L.C. CONDERMENT 1100 May Vol. 200	hington	STATE: UC COUNTRY: USA	ZIF: ZUUUS COMPUTER READABLE FORM:	MEDIUM TYPE: Floppy disk COMPUTER: IBM PC compatible	w	CURRENT APPLICATION DATA:	FILICATION WONDER. 03/00/00/1935B	PRIOR APPLICATION DATA: APPLICATION NUMBER: US 07/872,642	FILING DATE: 20-APR-1992 PRIOR APPLICATION DATA:	APPLICATION NUMBER: US 07/930,022 FILING DATE: 17-AHG-1992	_	۔ ت	OCKET NUMBE TION INFORM	TELEPHONE: (202)371-2571 TELEFAX: (202)371-2540	INFORMATION FOR SEQ ID NO: 3: SEQUENCE CHARACTERISTICS:	LENGTH: 653 amino acids TYPE: amino acid	TOPOLOGY: linear MOLECULE TYPE: protein	US-08-007-999B-3	Query Match 5.6%; Score 85.5; DB 2; Length 653;	21.2%; rative 22	SLEPMLDLAGDDTGAAVLTHLAPQLLEAG :: :	95 HAFTDAGVRLLMHPEQFIVLNSDRPEVRESSVRAMSAHARVMDGLGLARTPWN	DD 143 EFVSEALLVPEGCRFLHQERMDQCESSTRRHQEAQEACSSQGLILHGSG 191

```
dels 99; Gaps 10;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 LTH--LAPQLLEAG 94
| : | | 1
HPHHEVVPFHCLPG 142
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            AHARVMDGLGLARTPWN 147
: : | |
SQGLILHGSG----- 191
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        gth 653;
AND USES
                                            . C.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       Db 143 EFVSEA---
```

```
6
148 LLLIHGGKGGRGAELAALIPDLPDPVRLRLGLENDERAYSPAELLPICEATGTPLVFDAH 207
                                                                                                208 HHVVHDKLPDQEDPSVREWVLRARATWQPPEWQVVHLSNGIEGPQDRRHSHLIADFPSA- 266
                                                                                                                                          224 --------DPSTRSWPLGGRA-------EGGEDEEE---VESFPQPV 252
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               137 FYAVNSGQISIRHGMKGPNGVVVSDQAGGLDALAQARRLVRKGTPLIVCGAVDASICPWG 196
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       ----- RLLMHPEQFI ----- 112
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              77 LAAADWALADAGVEPESFDPLDMGVVTAGHAGGFEFGQGELQKLWAKGSQFVSAYQSFAW 136
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              18 GAADYWAATRGGKSGIGRITRFDPSSYP-ARLAGEIPGFEAAEHLPGRLLPQTDRVTRLS 76
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              48 GAADYCAAHD-----IRLYRLSSSLFPMLDLAGDDTGAAVLTHLAPQLL----- 91
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             ; OTHER INFORMATION: Description of Artificial Sequence: KS beta gene US-08-861-774E-27
                                                    -------PATPNPSGMAAG------
                                                                                                                                                                                                                                                                                                                                                                                                                                               APPLICANT: Waters, Barbara
APPLICANT: Miao, Vivian
APPLICANT: Miao, Vivian
APPLICANT: Ho, Yap
APPLICANT: Tong, Seow
TITLE OF INVENTION: METHOD FOR ISOLATION OF BIOSYNTHESIS GENES FOR TITLE OF INVENTION: BIOACTIVE MOLECULES
FILE REFERENCE: 9993-006
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  92;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        Query Match 5.5%; Score 84; DB 4; Length 403; Best Local Similarity 22.8%; Pred. No. 0.5; Matches 51; Conservative 17; Mismatches 64; Indels
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      148 -LLLLHGGKGGRGAELAALIPDLPDPVRLRLGLENDERAYSPAE 190
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     197 WVAQLAGGR------MSDSDEPARAYLPFDRDARGYLPGE 230
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           Sequence 12, Application US/08804227C
Patent No. 5876991
GENERAL INFORMATION:
APPLICANT: DeHoff, Bradley S.
APPLICANT: Ruhstoss, Stuart A.
APPLICANT: Rosteck, Paul R., Jr.
APPLICANT: Sutton, Kimberly L.
TITLE OF INVENTION: POLYKETIDE SYNTHASE GENES;
NUMBER OF SEQUENCES: 15
CORRESPONDENCE ADDRESS:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   113 -- VLNSDRPEVRES ---- SVRAMSAHARVMDGLGLAR ----
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       CURRENT APPLICATION NUMBER: US/08/861,774E CURRENT FILING DATE: 1997-05-22 NUMBER OF SEQ ID NOS: 94 SOFTWARE: Patentin Ver. 2.0
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         THOMAS G. PLANT 1501
                                                                                                                                                                                                                                                                                                                                                                         Sequence 27, Application US/08861774E Patent No. 6297007
                                               192 MLLPCGSDRFRGVEYVCCPP-----
                                                                                                                                                                                              267 ---YADVPQIEVEAKGKEE 282
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         ORGANISM: Artificial Sequence
                                                                                                                                                                                                                                                253 DDYFVEPPQAEEEEEEEE 271
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  92 --EAGHAFTDAGV------
                                                                                                                                                                                                                                                                                                                                                                                                                          GENERAL INFORMATION:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         ADDRESSEE:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         US-08-804-227C-12
                                                                                                                                                                                                                                                                                                                                                 US-08-861-774E-27
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        SEQ ID NO 27
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              Query Match
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    TYPE: PRT
                                               g
                                                                                              δλ
                                                                                                                                             Op
                                                                                                                                                                                                 ò
                                                                                                                                                                                                                                              qq
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                a
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                õ
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  q
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    ó
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                q
  ð
```

```
16;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    978 ADDAGIEHELRAVLPHLAAWHDRDRATARTAGLHYRVTWQAIEADAVRFS-----PSDRW 1032
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              768 RTKRLRVSHAFHSPRMDAMLAD------FRAVADTVDYHAPRLPVVS-----EVTG 812
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                76 DDTGAAVLTHLAPQLLEAGH-----112
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 813 DLADAA-----QLTDPGYWTRQVRQPVRFADAVRTASARDAATFIELGPDAVLCGMAE 865
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  113 -----VLNSDRPEVRESSVRAMSAHARVMDGLGLARTPWNLLLLHGGKGGRG 159
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               866 ESLAAEADVVFAPALRRGRPEGDTVLRAAASAYVR---GAGL---DW--AALYGGTGARR 917
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     -----AALIPDLPDP-VRLRLGLENDERAYSPAELLPICEATGTPL 202
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  203 VFDAH-HHVVHDKLP-----DQEDPSVREWVLRARATWQPPEWQVVHLSNGIEGPQDR- 254
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              28; Mismatches 126; Indels 114; Gaps
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                16 RTVTLSRYRALSPAEREAKLLDLYSSNIKTLRGAADYCAAHDIRLYRLSSSLFPMLDLAG 75
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        5.4%; Score 83; DB 2; Length 1891; 22.5%; Pred. No. 7.6;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    1033 LMVEHGOHTEC-----ADAAERALRAAGAE----VTRLVWPLE 1066
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     255 -----RHSHLIADFPSAYADVPQIEVEAKGKEEAIAALRLMAPFK 294
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               APPLICANT: Burgett, Stanley G.
APPLICANT: Rubstoss, Strart A.
APPLICANT: Rao, Nagaraja R.
APPLICANT: Rosteck, Paul R., Jr.
TITLE OF INVENTION: PLATENOLIDE SYNTHASE GENE
NUMBER OF SEQUENCES: 6
CORRESPONDENCE ADDRESS:
                                                                                                                                                                                                                  SOFTWARE: ASCI(DOS) Text only CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/804,227C FILING DATE: FEBRUARY 21, 1997
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          ADDRESSEE: PAUL R. CANTRELL 1138
STREET: LILLY CORPORATE CENTER
: LILLY CORPORATE CENTER INDIANAPOLIS
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           Sequence 6, Application US/08804198 Patent No. 5945320
                                                                                                                                                                                                                                                                                                                                                                                                                     REFERENCE/DOCKET NUMBER: X-(
TELECOMMUNICATION INFORMATION:
TELEPHONE: 317-276-2459
INFORMATION FOR SEQ ID NO: 12:
SEQUENCE CHARACTERISTICS:
                                                                                                                                                                                                                                                                                                                                                                                                   35,784
                                                                                                                                              MEDIUM TYPE: Floppy disk COMPUTER: IBM COMPATIBLE OPERATING SYSTEM: MS-DOS
                                                                                                                                                                                                                                                                                                                                            ATTORNEY/AGENT INFORMATION:
NAME: Plant, Thomas, G.
REGISTRATION NUMBER: 35,7
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    1891 amino acids
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    Best Local Similarity 22.5%
Matches 78; Conservative
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  , MOLECULE TYPE: peptide US-08-804-227C-12
                                                                                                                       COMPUTER READABLE FORM:
                                                                                                                                                                                                                                                                                                                            435
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     unknown
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              amino acid
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             GENERAL INFORMATION:
                                                                                                                                                                                                                                                                                                                       CLASSIFICATION:
                                                                           USA
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  160 AEL ----
                                                                                                  46285
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   TOPOLOGY:
                                              STATE: 1
COUNTRY:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       US-08-804-198-6
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    LENGTH:
  STREET:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               Query Match
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 q
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                δy
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     δλ
```

```
978 ADDAGIEHELRAVLPHLAAWHDRDRATARTAGLHYRVTWQAIEADAVRFS-----PSDRW 1032
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            813 DLADAA-----QLTDPGYWTRQVRQPVRFADAVRTASARDAATFIELGPDAVLCGMAE 865
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            113 ------VLNSDRPEVRESSVRAMSAHARVMDGLGLARTPWNLLLLHGGKGGRG 159
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         866 ESLAAEADVVFAPALRRGRPEGDTVLRAAASAYVR---GAGL---DW--AALYGGTGARR 917
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            160 AEL-----AALIPDLPDP-VRLRLGLENDERAYSPAELLPICEATGTPL 202
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            203 VFDAH-HHVVHDKLP-----DQEDPSVREWVLRARATWQPPEWQVVHLSNGIEGPQDR- 254
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         768 RTKRLRVSHAFHSPRMDAMLAD-----FRAVADTVDYHAPRLPVVS-----EVTG 812
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            76 DDTGAAVLTHLAPQLLEAGH----- 112
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         918 TDLPTYAFQHSRYWLAPASAAVAPATAAPSVRSVPEAEQDGALWAAVHAGDVASAAARLG 977
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              22.5%; Pred. No. 7.5; ative 28; Mismatches 126; Indels 114; Gaps
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            16 RTVTLSRYRALSPAEREAKLLDLYSSNIKTLRGAADYCAAHDIRLYRLSSSLFPMLDLAG 75
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     GENERAL INFORMATION:
APPLICANT: Katz, L
APPLICANT: Donadio, S
APPLICANT: Donadio, S
TITLE OF INVENTION: Recombinant DNA Method for Producing
TITLE OF INVENTION: Erythromycin Analogs
NUMBER OF SEQUENCES: 27
CORRESPONDENCE ADDRESS:
ADDRESSE: Edward H. Gorman
STREET: Abbott Laboratories D377/AP6D-2 one Abbott
STREET: Park Rd
CITY: Abbott Park
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      5.4%; Score 83; DB 2; Length 1891; 22.5%; Pred. No. 7.6;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  1033 LMVEHGQHTEC-----ADAAERALRAAGAE----VTRLVWPLE 1066
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  255 -----RHSHLIADFPSAYADVPQIEVEAKGKEEAIAALRLMAPFK 294
                                                                                                                                                COMPUTER: Macintosh 7.0 SPERATING SYSTEM: Macintosh 7.0 SOFTWARE: Microsoft Word 5.1 CURRENT APPLICATION DATA: US/08/804,198 FILING DATE:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          Sequence 4, Application US/07642734C Patent No. 5824513
                                                                                                                                                                                                                                                                                                                   ATTORNEY AGENT INFORMATION:
NAME: CANTRELL, PAUL R.
REGISTRATION NUMBER: 36,470
REFERENCE/DOCKET NUMBER: P911:
TELECOMMUNICATION INFORMATION:
TELEPHONE: 317-276-3885
INFORMATION FOR SEQ ID NO: 6:
                                                                                                                        E: Floppy disk
Macintosh
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              1891 amino acids
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           SEQUENCE CHARACTERISTICS:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         78; Conservative
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   rrPE: amino acid
; TOPOLOGY: unknown
; MOLECULE TYPE: peptide
US-08-804-198-6
                                                                                                  COMPUTER READABLE FORM:
                                                                                                                                                                                                                                                                                                   CLASSIFICATION: 435
CITY: INDIANAPOLIS
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 Best Local Similarity
                                             USA
                                                                                                                        MEDIUM TYPE:
                                                                        46285
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  US-07-642-734C-4
                                                COUNTRY:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                LENGTH:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         Query Match
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         Matches
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         QQ
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         qq
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               QΩ
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            Qγ
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               δλ
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               οy
```

```
13;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     1847 VVLAMRHGMLPRSLHADELSPHIDWESGAVEVLREEVPWPAGERPRRAGVSS----FGV 1901
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          74 AGDDTGAAVLTHLAPQLLEAGHAFTDAGVRLLMHPEQFIVLNSDRPE-VRESSVRAMSAH 132
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              133 ARVMDGLGLARTPWNL------LLLHGGKGGRGAELAALIPDLPDPVRLRLGLEN 181
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              182 DERAYSPAELLPICEATGTP-LVFDAHHHVVHDKLPDQEDPSVREWVLRARATWQPPE-- 238
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   83; Gaps
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            18 VTLSRYRALSP----AEREAKLLDLYSSNIKTLRGAADYCAAHDIRLYRLSSSLFPMLDL 73
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      1950 LRDTPELGLTDAAWTLATGRARFDVRAAVLGDDRAGVCAELDALAEGRP-----
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    5.4%; Score 82.5; DB 2; Length 3567; 23.8%; Pred. No. 24;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   Indels
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            APPLICANT: Donadio, S
APPLICANT: Kalz, L
APPLICANT: Mcalpine, J B
APPLICANT: Mcalpine, J B
TITLE OF INVENTION: Specific Polyketides
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            ADDRESSEE: Steven F. Weinstock
STREET: Abbott Laboratories D377/AP6D-2 One Abbott
STREET: Park Rd
                                                                                                                         SOFTWARE: Patentin Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/07/642,734C
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 28; Mismatches
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            2039 AESMSRCAEALSPHTDWKLLDVVRGDGGP 2067
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    -----WQVVHLSNGIEGP 251
                                                                                                                                                                                                                                                                                           REFERENCE/DOCKET NUMBER: 4952.US.01
TELECOMMUNICATION INFORMATION:
TELEPHONE: 708-937-9396
                                                                              COMPUTER: IBM PC compatible OPERATING SYSTEM: PC-DOS/MS-DOS
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             ; Sequence 4, Application US/08439009A
; Patent No. 6004787
                                                                                                                                                                                                                                                                         32652
                                                                                                                                                                                   FILING DATE: 17-JAN-91
CLASSIFICATION: 435
ATTORNEY/AGENT INFORMATION:
NAME: Danckers, Andreas M
REGISTRATION NUMBER: 32652
                                                                                                                                                                                                                                                                                                                                                                                INFORMATION FOR SEQ ID NO: 4:
                                                          Floppy disk
                                                                                                                                                                                                                                                                                                                                                                                                                         3567 amino acids
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 64; Conservative
                                                                                                                                                                                                                                                                                                                                                                                                         SEQUENCE CHARACTERISTICS:
                                                                                                                                                                                                                                                                                                                          TELEPHONE: 708-938-2623
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 , MOLECULE TYPE: protein US-07-642-734C-4
                                   COMPUTER READABLE FORM:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              CORRESPONDENCE ADDRESS:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        NUMBER OF SECUENCES:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  Abbott Park
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               Best Local Similarity
                                                                                                                                                                                                                                                                                                                                                                                                                                               amino acid
                                                                                                                                                                                                                                                                                                                                                                                                                                                                      linear
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          60064-3500
                 60064 - 3500
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              GENERAL INFORMATION:
  Sn
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          SO
                                                            MEDIUM TYPE:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                 TOPOLOGY:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                US-08-439-009A-4
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        COUNTRY:
COUNTRY:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           Query Match
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         RESULT 12
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 Matches
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                δ
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    qq
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 δ
```

COMPUTER READABLE FORM:

```
13;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      1950 LRDTPELGLTDAAWTLATGRARFDVRAAVLGDDRAGVCAELDALAEGRP------ 1998
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          1999 ----SADAVAPVISAPRKPVLVF-----PGQG----AQWVGMARDLLESSEVF 2038
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  182 DERAYSPAELLPICEATGTP-LVFDAHHHVVHDKLPDQEDPSVREWVLRARATWQPPE-- 238
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               74 AGDDTGAAVLTHLAPQLLEAGHAFTDAGVRLLMHPEQFIVLNSDRPE-VRESSVRAMSAH 132
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        133 ARVMDGLGLARTPWNL------LLLHGGKGGRGAELAALIPDLPDPVRLRLGLEN 181
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             Gaps
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        18 VTLSRYRALSP----AEREAKLLDLYSSNIKTLRGAADYCAAHDIRLYRLSSSLFPMLDL 73
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 APPLICANT: Turner, Jan R.
APPLICANT: Waldron, Clive
TITLE OF INVENTION: Biosynthetic Genes For Spinosyn Insecticide
TITLE OF INVENTION: Production
NUMBER OF SEQUENCES: 39
CORRESPONDENCE ADDRESS:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     Ouery Match 5.4%; Score 82.5; DB 3; Length 3567; Best Local Similarity 23.8%; Pred. No. 24; Matches 64; Conservative 28; Mismatches 94; Indels 83
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           ADDRESSEE: Dow AgroSciences LLC Patent Department STREET: 9330 Zionsville Road CITY: Indianapolis
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   2039 AESMSRCAEALSPHTDWKLLDVVRGDGGP 2067
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       -----WQVVHLSNGIEGP 251
                                                                                                            APPLICATION NUMBER: US/08/439,009A FILING DATE: 11-MAY-1995 CLASSIFICATION: 435
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               APPLICANT: Broughton, Mary C.
APPLICANT: Crawford, Kathryn P.
APPLICANT: Madduri, Krishnamurthy
APPLICANT: Merlo, Donald J.
APPLICANT: Treadware
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             Sequence 6, Application US/09036987A Patent No. 6143526
                                                                                                                                                                                                   NAME: Casuto, Dianne
REGISTRATION NUMBER: 40,943
REFERENCE/DOCKET NUMBER: 4952.
TELECOMMUNICATION INFORMATION:
TELEPHONE: 847-938-3137
INFORMATION FOR SEQ ID NO: 4:
SEQUENCE CHARACTERISTICS:
LENOTH: 3567 amino acids
TYPE: amino acids
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             COMPUTER READABLE FORM: MEDIUM TYPE: Floppy disk
                                                                                                                                                                                ATTORNEY/AGENT INFORMATION:
                                                                                                                                                                                                                                                                                                                                                                                                                           MOLECULE TYPE: protein US-08-439-009A-4
                                                                                                                                                                                                                                                                                                                                                                                                                  linear
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            GENERAL INFORMATION:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              Indiana
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          46268
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        US-09-036-987A-6
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                COUNTRY:
                                                                                                                                                                                                                                                                                                                                                                                                             TOPOLOGY:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              STATE:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          a
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               Ω
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         Q
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                q
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             ŏ
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       δλ
```

```
5311 RSLFQQAYAAGRILDGMDLVKVAAQLRPVFGSPGE-----LESLPKPVQLSRG---- 5358
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   183 ERAYSPAELLPICEATGTPLVFDAHHHVVHDKLPDQEDPSVREWVLRARATWQPPEWQVV 242
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          66 SLFPML--DLAGD-DTGAAVLTHLAPQLLEAGHAFTDAGVRLLMHPEQFIVLNSDRPEVR 122
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           123 ESSVRAMSAHARVMDGLGLARTPWNLLLLHGGKGGRGAELAALIPDLPDPVRLRLGLEND 182
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  Gaps
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       ----AADYCAAHDIRLYRLSS 65
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   APPLICANT: Waldron, Clive
TITLE OF INVENTION: Biosynthetic Genes For Spinosyn Insecticide
FILE REFERENCE: 50489 DIV1
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    DB 4; Length 5588;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       243 HLSNGIEGPQDRRHSHLIADFPSAYADVP---QIEVEAKGKEEAI 284
                            SOFTWARE: Patentin Release #1.0, Version #1.30 CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/09/036,987A
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               103;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               31; Mismatches
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      5.3%; Score 81.5;
21.8%; Pred. No. 64
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              CURRENT APPLICATION NUMBER: US/09/370,700 CURRENT FILING DATE: 1999-08-09 EARLIER APPLICATION NUMBER: US 09/36987 EARLIER FILING DATE: 1998-03-09 NUMBER OF SEQ ID NOS: 39 SOFTWARE: Patentin Ver. 2.0
                                                                   UMBER: US/09/036,987A
09-MAR-1998
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          ; ORGANISM: Saccharopolyspora spinosa
US-09-370-700-6
COMPUTER: IBM PC compatible OPERATING SYSTEM: PC-DOS/MS-DOS
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         22 RYRALSPAEREAKLLDLYSSNIKTLRG--
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         ; Sequence 6, Application US/09370700
; Patent No. 6274350
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          APPLICANT: Madduri, Krishnamurthy
                                                                                                                                         ATTORNEY/AGENT INFORMATION:
NAME: Stuart, Donald R
REGISTRATION NUMBER: 28,479
                                                                                                                                                                                                       REFERENCE/DOCKET NUMBER: 50
TELECOMMUNICATION INFORMATION:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        APPLICANT: Baltz, Richard H
APPLICANT: Broughton, Mary C
APPLICANT: Crawford, Kathryn P
                                                                                                                                                                                                                                                                                        INFORMATION FOR SEQ ID NO: 6: SEQUENCE CHARACTERISTICS:
                                                                                                                                                                                                                                                  : (317)337-4816
(317)337-4847
                                                                                                                                                                                                                                                                                                                                        5588 amino acids
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         APPLICANT: Treadway, Patti J
APPLICANT: Turner, Jan R
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  Conservative
                                                                                                                                                                                                                                                                                                                                                                                              MOLECULE TYPE: protein
                                                                                                                                                                                                                                                                                                                                                              amino acid
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          Best Local Similarity
                                                                                                                                                                                                                                                                                                                                                                                 linear
                                                                                                                      CLASSIFICATION:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       GENERAL INFORMATION:
                                                                                                    FILING DATE:
                                                                                                                                                                                                                                                  TELEPHONE:
                                                                                                                                                                                                                                                                                                                                                                                                                         US-09-036-987A-6
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   LENGTH: 5588
                                                                                                                                                                                                                                                                                                                                                                               TOPOLOGY:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               62;
                                                                                                                                                                                                                                                                          TELEFAX:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          9-004-046-60-SD
                                                                                                                                                                                                                                                                                                                                        LENGTH:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      Query Match
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               SEQ ID NO 6
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         TYPE: PRT
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      RESULT 14
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               Matches
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            δλ
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    qq
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             ά
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  q
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            ò
```

98; Gaps

67; Conservative 43; Mismatches 111; Indels

Matches

```
12;
                                                                                                                                                                                                                5263 ATGLRLPATLAFDHGTAAALAQHVCAQL---GTATAPAPRR-----TDDNDATEPV 5310
                                                                                                                                                                                                                                                                                    | : | |::||: | 5311 RSLEQQAYAAGRILDGMDLVKVAAQLRPVFGSPGE------LESLPKPVQLSRG---- 5358
                                                                                                                            5204 RLAGLSEGERROQVLQRVRADIAVVLGHGRSSDVDIEKPLAELGFDSLTAIELR-NRLAT 5262
                                                                                                                                                                                                                                                                                                                                                                                     66 SLFPML--DLAGD-DTGAAVLTHLAPQLLEAGHAFTDAGVRLLMHPEQFIVLNSDRPEVR 122
                                                                                                                                                                                                                                                         123 ESSVRAMSAHARVMDGLGLARTPWNLLLLHGGKGGRGAELAALIPDLPDPVRLRLGLEND 182
                                                                                                                                                                                                                                                                                                                                             183 ERAYSPAELLPICEATGTPLVFDAHHHVVHDKLPDQEDPSVREWVLRARATWQPPEWQVV 242
                                               Gaps
                                                                                      ----AADYCAAHDIRLYRLSS 65
                                             89;
    DB 4; Length 5588;
Match 5.3%; Score 81.5; DB 4; Length 55 Local Similarity 21.8%; Pred. No. 64; es 62; Conservative 31; Mismatches 103; Indels
                                                                                                                                                                                                                                                                                                                                                                                                                                                         243 HLSNGIEGPQDRRHSHLIADFPSAYADVP---QIEVEAKGKEEAI 284
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        GENERAL INFORMATION:
APPLICANT: Hamel, Josee
APPLICANT: Brodeur, Bernard R
APPLICANT: Martin, Denis
TITLE OF INVENTION: HEAT SHOCK PROTEIN HSP72 FROM
TITLE OF INVENTION: STREPTOCOCCUS PNEUMONIAE
NUMBER OF SEQUENCES: 6
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   SOFTWARE: Patentin Release #1.0, Version #1.30 CURRENT APPLICATION DATA:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         CORRESPONDENCE ADDRESS:
ADDRESSEE: Fish & Neave
STREET: 1251 Avenue of the Americas
                                                                                 22 RYRALSPAEREAKLLDLYSSNIKTLRG-----
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        United States of America
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                APPLICATION NUMBER: US/08/472,534
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       RESULT 15
US-08-472-534-5
Sequence 5, Application US/08472534
; Patent No. 5919620
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        27,794
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        REFERENCE/DOCKET NUMBER: Bir
TELECOMMUNICATION INFORMATION:
TELEPHONE: 212-596-9000
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     TELERAX: 212-596-9090
TELEX: 14-8367
INFORMATION FOR SEQ ID NO: 5:
SEQUENCE CHARACTERISTICS:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        Floppy disk
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           ATTORNEY/AGENT INFORMATION:
NAME: Haley Jr, James F
REGISTRATION NUMBER: 27,7
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              : 607 amino acids
amino acid
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      ; MOLECULE TYPE: protein US-08-472-534-5
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               COMPUTER READABLE FORM: MEDIUM TYPE: Floppy
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         CLASSIFICATION: 424
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         linear
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             STREET: 122.
CITY: New York
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  New York
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    FILING DATE:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      COUNTRY: U
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    TOPOLOGY:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           LENGTH:
  Query Match
                                           Matches
                                                                                                                            qq
                                                                                                                                                                       οy
                                                                                                                                                                                                                                                                                               Д
                                                                                                                                                                                                                                                                                                                                                                                   q
                                                                                                                                                                                                                                                                                                                                                                                                                                 δλ
                                                                                                                                                                                                                                                                                                                                                                                                                                                                       q
                                                                                                                                                                                                                                                         ò
                                                                                                                                                                                                                                                                                                                                             Óγ
```

Score 81; DB 2; Length 607; Pred. No. 2.1;

5.3%;

Query Match Best Local Similarity

```
| :: | :| | : | 111 | 111 | 111 | 111 | 111 | 111 | 111 | 111 | 111 | 111 | 111 | 111 | 111 | 111 | 111 | 111 | 111 | 111 | 111 | 111 | 111 | 111 | 111 | 111 | 111 | 111 | 111 | 111 | 111 | 111 | 111 | 111 | 111 | 111 | 111 | 111 | 111 | 111 | 111 | 111 | 111 | 111 | 111 | 111 | 111 | 111 | 111 | 111 | 111 | 111 | 111 | 111 | 111 | 111 | 111 | 111 | 111 | 111 | 111 | 111 | 111 | 111 | 111 | 111 | 111 | 111 | 111 | 111 | 111 | 111 | 111 | 111 | 111 | 111 | 111 | 111 | 111 | 111 | 111 | 111 | 111 | 111 | 111 | 111 | 111 | 111 | 111 | 111 | 111 | 111 | 111 | 111 | 111 | 111 | 111 | 111 | 111 | 111 | 111 | 111 | 111 | 111 | 111 | 111 | 111 | 111 | 111 | 111 | 111 | 111 | 111 | 111 | 111 | 111 | 111 | 111 | 111 | 111 | 111 | 111 | 111 | 111 | 111 | 111 | 111 | 111 | 111 | 111 | 111 | 111 | 111 | 111 | 111 | 111 | 111 | 111 | 111 | 111 | 111 | 111 | 111 | 111 | 111 | 111 | 111 | 111 | 111 | 111 | 111 | 111 | 111 | 111 | 111 | 111 | 111 | 111 | 111 | 111 | 111 | 111 | 111 | 111 | 111 | 111 | 111 | 111 | 111 | 111 | 111 | 111 | 111 | 111 | 111 | 111 | 111 | 111 | 111 | 111 | 111 | 111 | 111 | 111 | 111 | 111 | 111 | 111 | 111 | 111 | 111 | 111 | 111 | 111 | 111 | 111 | 111 | 111 | 111 | 111 | 111 | 111 | 111 | 111 | 111 | 111 | 111 | 111 | 111 | 111 | 111 | 111 | 111 | 111 | 111 | 111 | 111 | 111 | 111 | 111 | 111 | 111 | 111 | 111 | 111 | 111 | 111 | 111 | 111 | 111 | 111 | 111 | 111 | 111 | 111 | 111 | 111 | 111 | 111 | 111 | 111 | 111 | 111 | 111 | 111 | 111 | 111 | 111 | 111 | 111 | 111 | 111 | 111 | 111 | 111 | 111 | 111 | 111 | 111 | 111 | 111 | 111 | 111 | 111 | 111 | 111 | 111 | 111 | 111 | 111 | 111 | 111 | 111 | 111 | 111 | 111 | 111 | 111 | 111 | 111 | 111 | 111 | 111 | 111 | 111 | 111 | 111 | 111 | 111 | 111 | 111 | 111 | 111 | 111 | 111 | 111 | 111 | 111 | 111 | 111 | 111 | 111 | 111 | 111 | 111 | 111 | 111 | 111 | 111 | 111 | 111 | 111 | 111 | 111 | 111 | 111 | 111 | 111 | 111 | 111 | 111 | 111 | 111 | 111 | 111 | 111 | 111 | 111 | 111 | 111 | 111 | 111 | 111 | 111 | 111 | 111 | 111 | 111 | 111 | 111 | 11
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                369 LGIETMGGVFTKLIDRNTTIPTSKSQVFSTAADNQPA--VDIHVLQG-ERPMAADNKTLG 425
                                                                                                                                                                                                                                                                                                                                                                                                80 AAVLTHLAPQ------AFTDAGVRLL 105
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                204 OKIIDHLVAEFKKENGIDI.STDKMAMORLKDAAEKAKKDLSGVTSTQISLPFITAGEAGP 263
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          106 MHPE-----QFTVLNSDRPEVRESSVRAMSAHARVMDGLGLARTPWNLLLLHGGKGGRG 159
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           264 LHLEMTLTRAKFDDLTRDLVERTKVPVR-----QALSDAGLSLSEIDEVILVGG----- 312
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           160 AELAALIPOLPOPVRLRLGLENDERAYSPAELLPICEA-TGTPLVFDAHHHVVHD---- 213
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           313 ---STRIFAVVEAVKAETGKE-PNKSVNPDEVVAMGAAIQGGVITGDVKDVVLLDVTPLS 368
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                214 ------KLPDQED--PSVREWVLRARATWQPPEWQVVHLSNGIEGPQDRRHSHL- 259
26 LSPAEREAKLLDLYSSNIKTLRGAADYCAAHDIRLYRLSSSLFPMLD-----LAGDDTG 79
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      Search completed: January 15, 2002, 13:51:29
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                260 --- IADFFSAYADVPQLEV 275
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      qq
                                                                                                                                                                                                                                                                                                                                                                                                ŏ
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          δy
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                qq
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                qq
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           q
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           ζý
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     à
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      δy
```

```
Compugen Ltd.
GenCore version
Copyright (c) 1993 - 2000
```

OM protein - protein search, using sw model

January 15, 2002, 14:07:20; Search time 45.07 Seconds (Without alignments) 471.835 Million cell updates/sec

US-09-724-296-38 Perfect score: Title:

1 MGTTGLESLSLGDRGAAPTV......HKLNKYHDWLWENHEKLSLS 580 Sequence:

BLOSUM62 Scoring table: 100059 segs, 36664827 residues Searched:

Gapop 10.0 , Gapext 0.5

100059 Total number of hits satisfying chosen parameters:

Minimum DB seq length: 0 Maximum DB seq length: 20000000000 .

Maximum Match 100% Post-processing: Minimum Match 0%

Listing first 45 summaries

SwissProt_39:* Database : Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

		d			SUMMARIES		
Result		Query					
NO.	Score	Match	Length	DB	ID	Description	tion
1	208.5	6.8	486	-	MEC2_HUMAN	P51608	homo sapien
7	201	9.9	492	7	MEC2_RAT	000566	rattus norv
m	198.5	6.5	484	٦	MEC2_MOUSE	09z2d6	mus musculu
4	145	4.7	1301	П	SAC3_YEAST	P46674	saccharomyc
2	145	4.7	3418	7	BRC2_HUMAN	P51587	homo sapien
9	138.5	4.5	200	Н	GAR2_SCHPO	P41891	schizosacch
7	137	4.5	260	П	YJK5_YEAST	P42948	saccharomyc
œ	136.5	4.5	598	7	CYL1_HUMAN	P35663	homo sapien
6	133	4.4	1447	Н	BUD4_YEAST	P47136	saccharomyc
10	131.5	4.3	997	-	BIR1_SCHPO	014064	schizosacch
11	122	4.0	3924	٦	ANK2_HUMAN	001484	homo sapien
12	120.5	3.9	1147	Н	AC15_HUMAN	P35251	homo sapien
13	120.5	3.9	1233	7	YFI6_YEAST	P43597	saccharomyc
14	120	3.9	1131	П	AC15_MOUSE	P35601	mus musculu
15	120	3.9	1658	7	YM67_YEAST	003661	saccharomyc
16	119	9.6	886	-	ORC1_KLULA	P54788	kluyveromyc
17	118	9.	704		NP14_RAT	P41777	rattus norv
18	118	σ, σ,	3329	_	BRC2_MOUSE	P97929	mus musculu
19	117	3.8	1273		YAR2_SCHPO	010135	schizosacch
50	117	3.8	1850	П	VIT2_CHICK		gallus gall
21	116.5	3.8	633	7	MLH_TETTH		tetrahymena
22	116	3.8	1290	7	XCPC_XENLA	P50532	xenopus lae
23	116	3.8	3256	7	K167_HUMAN		homo sapien
24	115.5	3.8	2156	Н	ORP1_HUMAN	ıo	homo sapien
52	114.5	3.7	970	1	PSU1_YEAST	P53550	saccharomyc
56	114.5	3.7	1213	_	FMN_CHICK		gallus gall
27	114	3.7	783	-	YAYB_SCHPO		schizosacch
28	114	3.7	802	Н	TAC1_HUMAN	075410	homo sapien
29	114		1539	-	Y373_HUMAN	015078	homo sapien
30	113.5		1164		BAG_STRAG	P27951	streptococc
31	113	3.7	482	7	YSR2_CAEEL	056600	caenorhabdi
32	113	3.7	727	7	PEC1_MOUSE	008481	mus musculu
33	112.5	3.7	919	-	HS7C_TRYBB		trypanosoma

Q09750 schizosacch	P30414 homo sapien	Q01448 saccharomyc	P49452 mus musculu	P35201 saccharomyc	Q24595 drosophila	P20357 mus musculu	P35662 bos taurus	P39526 saccharomyc	P17141 mus musculu	Q13435 homo sapien	P20676 saccharomyc
PST1_SCHPO	NKCR_HUMAN	HPC2_YEAST	CENC_MOUSE	MlF2_YEAST	XPC_DROME	MAP2_MOUSE	CYL1_BOVIN	YJU7_YEAST	ZF37_MOUSE	S3B2_HUMAN	NUP1_YEAST
-	~	,		-	Н	-	-	-	-	-	Н
1522	1462	623	906	549	1293	1828	299	2014	594	872	1076
3.7	3.7	3.6	3.6	3.6	3.6	3.6	3.6	3.6	3.6	3.6	3.6
_	12	1.5	11.5	111	111	111	110.5	110.5	110	110	110
112.5	-	Ξ	H								

ALIGNMENTS

```
TISSUB-skeletal muscle;
MEDLINE-96327611; PubMed=8672133;
D'Esposito M., Quaderi N.A., Ciccodicola A., Bruni P., Esposito T.,
D'Esposito M., Brown S.D.M.;
D'Urso M., Brown S.D.M.;
"Isolation, physical mapping, and Northern analysis of the X-linked human gene encoding methyl CpG-binding protein, MECP2.";
Mamm. Genome 7:533-535(1996).
                                                                                                                                                                                                                                                                                                                                                                                                                                              2 (MECP2) to
                                                                                                                               Homo sapiens (Human).
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    "Rett syndrome is caused by mutations in x-linked MECP2, encoding methyl-CpG-binding protein 2.";
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                Amir R.E., Van den Veyver I.B., Wan M., Tran C.Q., Francke U., Zoghbi H.Y.;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              SEQUENCE OF 10-486 FROM N.A.

Bechimaid K., Bauer D., Brenner V., Drescher B., Coy J.,

Kioschis P., Korn B., Nyakatura G., Platzer M., Poustka A.,

Sandoval N., Rosenthal A.;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   Reichwald K., Rosenthal A., Kloschis P., Platzer M.; "Mapping and sequence analysis of the human MeCP2 locus."; Submitted (OCT-1997) to the EMBL/GenBank/DDBJ databases.
                                                                                                                                                                                                                                                                                                              TISSUE=Placenta;
Thiesen J., Straetling W.H.;
Submitted (APR-1997) to the EMBL/GenBank/DDBJ databases
                                                                                                                                                                                                                                                                                                                                                                                       SEQUENCE FROM N.A.
MEDLINE=97130625; PubMed=8976388;
Vilain A., Apiou F., Vogt N., Dutrillaux B., Malfoy B.;
"Assignment of the gene for methyl-CpG-binding protein 2 human chromosome band Xq28 by in situ hybridization.";
Cytogenet. Cell Genet. 74:293-294(1996).
                                                                                                                                                                                                                                        Kudo S., Fukuda M.;
Submitted (SEP-1995) to the EMBL/GenBank/DDBJ databases.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        Submitted (DEC-1996) to the EMBL/GenBank/DDBJ databases
                 P51608; O15233;
01-CCT-1996 (Rel. 34, Created)
01-CCT-1996 (Rel. 34, Last sequence update)
02-GCT-1090 (Rel. 40, Last annotation update)
METHYL-CPG-BINDING PROTEIN 2 (MBCP-2 PROTEIN) (MECP2).
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            VARIANTS RTT TRP-106; CYS-133; SER-155 AND MET-158
   486 AA
   PRT;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               MEDLINE=99438392; PubMed=10508514;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              SEQUENCE OF 10-486 FROM N.A.
   STANDARD;
                                                                                                                                                                                                                                                                                                 SEQUENCE FROM N.A.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    SEQUENCE FROM N.A.
                                                                                                                                                                                                                          SEQUENCE FROM N.A.
                                                                                                                                                                                    NCBI_TaxID=9606;
MEC2_HUMAN
```

```
This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation the Buropean Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (see http://www.isb-sib.ch/announce/or send an email to license@isb-sib.ch).
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 -!- TISSUE SPECIFICITY: PRESENT IN ALL ADULT SOMATIC TISSUES TESTED.
-!- TISSUE SPECIFICITY: PRESENT IN ALL ADULT SOMATIC TISSUES TESTED.
-!- DISEASE: DEPECTS IN MECP2 ARE THE CAUSE OF RETT SYNDROME (RTT), AN
X-LINKED DOMINANT DISCASE. RTI IS A PROGRESSIVE NEUROLOGIC
BETARDATION IN FEMALES. PATIENTS APPEAR TO DEVELOP NORMALLY UNTIL
6 TO 18 MONTHS OF AGE, THEN GRADUALLY LOSE SPEECH AND PURPOSEFUL
HAND MOVEMENTS AND DEVELOP MICROCEPHALY. SETZURES, AUTISM, ATAXIA,
INTERMITTAL RESENSION, THE CONDITION STABILIZES AND PATIENTS
AFTER INTITAL RESENSION, THE CONDITION STABILIZES AND PATIENTS
USUALLY SURVIVE INTO ADULTHOOD
USUALLY SURVIVE INTO ADULTHOOD
-: SIMILARITY: CONTAINS (R)GRP(K) MOTIFS AND SPEK MOTIFS THAT HAVE
BEEN FOUND TO BIND TO MINOR GROOVE OF AT-RICH DNA.
-: SIMILARITY: CONTAINS 1 METHYL-BINDING DOMAIN (MBD).
                                                                                                                                                                                                      VARIANTS RTT W-106; F-124; C-13; C-134; R-152; M-158 AND C-306.
MEDLINE=20439334; PubMed=10991688;
Obata K., Matsuishi T., Yamashita Y., Fukuda T., Kuwajima K.,
Horiuchi I., Naqamitsu S., Iwanaga R., Kimura A., Omori I., Endo S.,
Mori K., Kondo I.;
"Mutation analysis of the methyl-CpG binding protein 2 gene (MECP2) in
patients with Rett syndrome."
J. Med. Genet. 37:608-610(2000).
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      -i- FUNCTION: CHROMOSOMAL PROTEIN THAT BINDS TO METHYLATED DNA. IT CAN BIND SPECIFICALLY TO A SINGLE METHYL-CPG PAIR. IT IS NOT INFLUENCED BY SEQUENCES FLANKING THE METHYL-CPGS. MEDIATES TRANSCRIPTIONAL REPRESSION THROUGH INTERACTION WITH HISTONE DEACETYLASE AND THE COREPRESSOR SINAA.
                                                           MEDLINE-20465115; PubMed-11007980; Ozritoc A., Lam C., Galli L., Dotti M.T., Hayek G., Tong S.F., Poon P.M., Zappella M., Federico A., Sorrentino V.; "MECP2 mutation in male patients with non-specific X-linked mental
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             Transcription regulation; Repressor; DNA-binding; Nuclear protein;
                                                                                                                                                                                                                                                                                                                                                                                               VARIANTS RTT R-101; W-106; M-158 AND C-306, AND VARIANT K-397. MEDILME-20439355; Pubmed-10991689; Medilme-20439355; Pubmed-10991689; Medilme X., Woods C.G., Latlif F., Webb T.; "Mutations in the MECP2 gene in a cohort of girls with Rett
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        POLY-ALA. POLY-HIS.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                MBD.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  Med. Genet. 37:610-612(2000).
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    EMBL, Y12643; CAA73190.1; -. EMBL, X99686; CAA68001.1; -. EMBL, AF030876; AAC08757.1; -. EMBL, AF031078; AAC08758.1; -.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 Disease mutation; Polymorphism DOMAIN 96 149 ME
Nat. Genet. 23:185-188(1999).
                                                                                                                                                                   FEBS Lett. 481:285-288(2000).
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     InterPro; IPR000637; AT_hook
InterPro; IPR001739; MBD.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 EMBL; L37298; AAC32737.1; -.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    X89430; CAA61599.1; -. X94628; CAA64331.1; -.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         Pfam; PF01429; MBD; 1.
SMART; SM00384; AT_hook; 1.
SMART; SM00391; MBD; 1.
                                         VARIANT RTT VAL-140.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            366
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              MIM; 300005; -.
MIM; 312750; -.
                                                                                                                                               retardation.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     syndrome.";
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        GENOME
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        DOMAIN
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      EMBL;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            EMBL;
```

```
17;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            Indels 129; Gaps
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           128 LHKNGETSLKPEDFDFTVLSKRGIKSRYKDCSMAALTSHLQNQSNNSNWNLRTRSKCKKD 187
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              188 VFMPPSSSSELQESRGL---SNFTSTHLLLKEDEGVDDVNFRKVRKPKGKVTILKGIPIK 244
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         176 ----PKSFKAPGTGRGRGRPKGSGTTRPKAATSEGVQVK--RVLEKSPGK--LLVKMPF- 226
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  245 KTKKGCRKSCSGFVQSDS----KRESVCNKADAESEPVAQKSQLDRTVCISDAGACGETL 300
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        227 QTSPGGKAEGGGATTSTQVMVIKRPGRKRKAEADPQAIPKKR------GRKPGSVV 276
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      277 AAAAAEAKKKAVKESSIR-----SVQETVLPIKK-----RKTRET----V 312
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               361 KVEVVERKEHLHTDIL--KRGSEMDNNCSPTRKDFTGEKIFQEDTIPRTQIERRKTSLYF 418
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    354
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             81 ECRKSV------PCGWERVVKQRLFGKTAGRFDVYFISPQGLKFRSKSSLANY 127
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  ---ATAGT 80
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       22 KDKPLKFKKVKKDKKEEKEGKHEPVQPSAHHSAEPAEAGKAETSEGSGSAPAVPEASASP 81
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                301 SVTSEENSLVKKKERSLSSGSNFCSEQKTSGIINKFCSAKDSEHNEKYEDTFLESEEIGT
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    313 SIEVKEVVKPLLVSTLGEKSGKGLKTCKSPGRK-----SKESSPKGR-----
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   6.8%; Score 208.5; DB 1; Length 486; 23.9%; Pred. No. 2.5e-06;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  -> G (IN REF. 3).
EB6A33233AEDA566 CRC64;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             PAVP -> RLC (IN REF. 5).
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            Matches 104; Conservative 53; Mismatches 149;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  37 EDVAMELERVGEDEEQMMIKRSSECNPLLQEPIASAQFG---
                                                                             R -> W (IN RTT).
/FTId=VAR_010272.
                                                                                                                                                                                                                      _010273.
                                                                                                                                                                                                                                                       S -> C (IN RTT).
/FTId=VAR_010278.
                                                                                                                                                                                                                                                                                                                                                                                                                        F -> S (IN RTT).
/FTId=VAR_010274.
T -> M (IN RTT).
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               R -> C (IN RTT).
/FTId=VAR_010282.
                       P -> R (IN RTT).
/FTId=VAR_010276.
                                                                                                                                                                                                                                                                                                                                    _010279.
                                                                                                                                                                                                                                                                                                                                                                  P -> R (IN RTT). /FTId=VAR_010280.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              /FTId=VAR_010281.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               /FTId=VAR_010283.
                                                                                                                                      L -> F (IN RTT).
/FTId=VAR_010277
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              /FTId=VAR_010275
                                                                                                                                                                                                                                                                                                                A -> V (IN RTT').
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              01-APR-1993 (Rel. 25, Created)
01-APR-1993 (Rel. 25, Last sequence update)
                                                                                                                                                                                                 R -> C (IN
                                                                                                                                                                                                                         /FTId=VAR
                                                                                                                                                                                                                                                                                                                                          /FTId=VAR_
  POLY - PRO.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            PRT;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        ы
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   52440 MW;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              Best Local Similarity 23.9%;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             419 SSKYNKEALSPPRRK 433
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                355 ----SSSASSPPKKE 365
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            STANDARD;
                                                                                                                                                                                                                                                    134
                                                                                                                                                                                                                                                                                                             1.40
                                                                                                                                                                                                                                                                                                                                                                                                                              155
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         201
                                                                                                                                         124
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     290
486 AA;
                                                                                                                                                                                                                                                                                                          140
                                                                                106
                                                                                                                                         124
                                                                                                                                                                                              133
                                                                                                                                                                                                                                                       134
                                                                                                                                                                                                                                                                                                                                                                                                                           155
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 158
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         201
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               306
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     397
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         MEC2_RAT
000566;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      Query Match
DOMAIN
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                SEQUENCE
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             CONFLICT
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        CONFLICT
                                                                                   VARIANT
                                                                                                                                         VARIANT
                                                                                                                                                                                                 ARIANT
                                                                                                                                                                                                                                                       VARIANT
                                                                                                                                                                                                                                                                                                                VARIANT
                                                                                                                                                                                                                                                                                                                                                                     VARIANT
                                                                                                                                                                                                                                                                                                                                                                                                                           VARIANT
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 VARIANT
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         VARIANT
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  VARIANT
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        VARIANT
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   2
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 MEC2_RAT
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   RESULT
  SO SET THE SET
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            q
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        ŏ
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 ŏ
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         q
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  δŏ
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     qq
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                ŏ
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      Db
```

```
This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation-the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See http://www.isb-sib.ch/announce/or send an email to license@isb-sib.ch).
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 18;
                                                                                                                                                                                                                                 Furification, sequence, and cellular localization of a novel chromosomal protein that binds to methylated DNA.";

cell 69:905-914(1992).

-!- FUNCTION: CHROMOSOMAL PROTEIN THAT BINDS TO METHYLATED DNA. IT CAN BIND SPECIFICALLY TO A SINGLE METHYL-CPG PATR. IT IS NOT INFLUENCED BY SEQUENCES FLANKING THE METHYL-CPG. MEDIATES TRANSCRIPTIONAL REPRESSION THROUGH INPERACTION WITH HISTONE DEACETYLASE AND THE CORPRESSOR SIN3A (BY SIMILARITY).

-!- SUBCELLULAR LOCATION: NUCLEAR. COLOCALIZED WITH METHYL-CPG IN THE
                                                                          Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.
NCBL_TaxID=10116;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 ---PCGWERVVKQRLFGKTAGRFDVYFISP 113
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    68 SGSAPAVPEASASPKORRSIIRDRGPMYDDPTLPEGWTRKLKORKSGRSAGKYDVYLINP 127
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         114 QGLKFRSKSSLANYLHKNGETSLKPEDFDFTVLSKRGIKSRYKDCSMAALTSHLQNQSNN 173
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                174 SNWNLRTRSKCKKDVFMPPSSSSELQESRGLSNFTSTHLLLKEDEGVDDVNFRKV-RKPK 232
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           Gaps
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  34 LRKE------DVAMELERVGEDEEQMMIKRSSECNPLLQEPIASAQFG----- 75
                                                                                                                                                                                                                                                                                                                                                                                                                                                         -!- TISSUE SPECIFICITY: PRESENT IN ALL ADULT SOMATIC TISSUES TESTED.
-!- SIMILARITY: CONTAINS (R)GRP(K) MOTIFS AND SPKK MOTIFS THAT HAVE BEEN FOUND TO BIND TO MINOR GROOVE OF AT-RICH DNA.
-!- SIMILARITY: CONTAINS 1 METHYL-BINDING DOMAIN (MBD).
                                                                                                                                                                                                                         Lewis J.D., Meehan R.R., Henzel W.J., Maurer-Fogy I., Jeppesen P.,
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          Repressor; DNA-binding; Nuclear protein.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     8 LRKEKSEDODLOGLKEKPLKFKKVKKDKKEDKEGKHEPLQPSAHHSAEPAGKAETSES
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         -----REOKPPKK-----PKSPKAPGTGRGRPKGSGTGRPKAAASEGVOVKRVLEKSP
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     233 GKVTILKGIPIKKTKKGCRKSCSGFVQSDS----KRESVCNKADAESEPVAQKSQLDRT
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            Indels 150;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     6.6%; Score 201; DB 1; Length 492; 22.9%; Pred. No. 7.8e-06;
20-AUG-2001 (Rel. 40, Last annotation update)
METHYL-CPG-BINDING PROTEIN 2 (MECP-2 PROTEIN) (MECP2).
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         A67E705C68BA2D38 CRC64;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          57; Mismatches 156;
                                                                                                                                                           SEQUENCE FROM N.A., AND PARTIAL SEQUENCE.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    POLY-PRO.
POLY-THR.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 POLY-HIS.
                                                                                                                                                                                                MEDLINE=92298389; PubMed=1606614;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           53047 MW;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        EMBL; M94064; AAA41584.1; -.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              ----ATAGTECRKSV---
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       ranscription regulation;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            Conservative
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               InterPro; IPR001739; MBD.
                                                            Rattus norvegicus (Rat).
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              Pfam; PF01429; MBD; 1.
SMART; SM00391; MBD; 1.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               PIR; A41907; A41907
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       492 AA;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      Best Local Similarity
Matches 108; Conserv
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         SEQUENCE
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     Query Match
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 DOMAIN
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    DOMAIN
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        DOMAIN
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               168
  Db
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  a
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            q
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           a
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        ŏ
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   δý
```

```
between the Swiss Institute of Bioinformatics and the EMBL outstation—
the European Bioinformatics Institute. There are no restrictions on its
use by non-profit institutions as long as its content is in no way
modified and this statement is not removed. Usage by and for commercial
entities requires a license agreement (See http://www.isb-sib.ch/announce/
or send an email to license@isb-sib.ch).
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     This SWISS-PROT entry is copyright. It is produced through a collaboration
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   Reichwald K., Thiessen J., Wiehe T., Kioschis P., Straetling W.H., Rosenthal A., Platzer M.; "Comparative analysis of the methyl CpG binding protein 2 locus in man
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     CAN
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 and mouse reveals new untranslated sequences.;
Submitted (JUN-1999) to the EMBL/GenBank/DbbJ databases.
-!-FUNCTION: CHROMOSOMAL PROTEIN THAT BINDS TO METHYLATED DNA. IT CAN BIND SPECIFICALLY TO A SINGLE METHYL-CPG PAIR. IT IS NOT INFLUENCED BY SEQUENCES FLANKING THE METHYL-CPG PAIR. IT IS NOT TRANSCRIPTIONAL REPRESSION THROUGH INTERACTION WITH HISTONE DEACETYLASE AND THE COMPRESSOR SINJA (BY SIMILARITY).
-!- SUBCELLULAR LOCATION: NUCLEAR. COLOCALIZED WITH METHYL-CPG IN THE
                                             307
                                                                                 348 YEDTFLESEEIGTKVEVVERKEHLHTDIL--KRGSEMDNNCSPTRKDFTGEKIFQEDTIP 405
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     Coy J.F., Sedlacek 2., Beachner D., Delius H., Poustka A.;
"A complex pattern of evolutionary conservation and alternative polyadenylation within the long 3'-untranslated region of the methyl-cpG-binding protein 2 gene (MeCP2) suggests a regulatory role in gene
288 VCISDAGACGETLSVTSEENSLVKKKERSLSSGSNFCSEQKTSGIINKFCSAKDSEHNEK 347
                                                                                                                          308 TRET-----VSIEVKEVVKPLLVSTLGEKSGKGLKTCKSPGRK-----SKESSP 351
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        Hendrich B., Bird A.;
"Identification and characterization of a family of mammalian methyl-
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    -!- SIMILARITY: CONTAINS (R)GRP(K) MOTIFS AND SPKK MOTIFS THAT HAVE
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
                                        269 -----GRKPGSVVAAAAAEAKKKAVKESSIR-----SVQETVLPIKK------RK
                                                                                                                                                                              444
                                                                                                                                                                                                                352 КGR-----SSSASSPPKKEHHHHHAESPKAPMPLLPPPPP 388
                                                                                                                                                                       406 RTQIERRKTSLYFSSKYNKEALSPPRR------KAFKKWTPPRSP
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        BEEN FOUND TO BIND TO MINOR GROOVE OF AT-RICH DNA. -!- SIMILARITY: CONTAINS 1 METHYL-BINDING DOMAIN (MBD).
                                                                                                                                                                                                                                                                                                                                                                 20-AUG-2001 (Rel. 40, Created)
20-AUG-2001 (Rel. 40, Last sequence update)
20-AUG-2001 (Rel. 40, Last annotation update)
METHYL-CPG-BINDING PROTEIN 2 (MECP-2 PROTEIN) (MECP2).
                                                                                                                                                                                                                                                                                                                             484 AA
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   CpG binding proteins."; Mol. Cell. Biol. 18:6538-6547(1998).
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        Hum. Mol. Genet. 8:1253-1262(1999).
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 SEQUENCE FROM N.A. MEDLINE=99299240; Pubmed=10369871;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             STRAIN=C57BL/6;
MEDLINE=98449942; PubMed=9774669;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             EMBL, AF072251; AAC68880.1; -.
EMBL, A7133922; CAB46495.1; -
EMBL, AF121351; AAF22116.1; -.
EMBL, AF1S6181; AAF33024.1; -.
                                                                                                                                                                                                                                                                                                                             STANDARD;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                              Mus musculus (Mouse).
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           SEQUENCE FROM N.A.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    SEQUENCE FROM N.A.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                NCBI_TaxID=10090;
                                                                                                                                                                                                                                                                                                                             MEC2_MOUSE
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      expression.
                                                                                                                                                                                                                                                                                                                                                    Q9Z2D6;
                                                                                                                                                                                                                                                                                  RESULT 3
MEC2_MOUSE
                                                                                                                                                                                                                                                                                                                             a
                                                                                                                            qq
                                                                                   δ
```

```
AC AC
 δ
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      q
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  δŏ
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        qq
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      δŏ
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              q
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        ŏ
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 qq
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            \tilde{\alpha}
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                ŏ
                                                                                                                                                                                                                          11;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                              300 VLPIKKRKTRETVSIEVKEVVKPLLVSTLGEKSGKGLKTCKSPGRKSKESSPKGRSSSAS 359
                                                                                                                                                                                                                                                                                                                 183 KCKKDVFMPPSSSSELQESRGLSNFTSTHLLLKEDEGVDDVNFRKVRKPKGK-----VT 236
                                                                                                                                                                                                                                                                                                                                                                                                           237 ILKGIPIKKTKKGCRKSCSGFVQSDSKRE----SVCNKADAESEPVAQKS----QLDRT 287
                                                                                                                                                                                                                                                                                                                                                                                                                                                               288 VCISDAGACGETLSVTSEE-----NSLVKKKERSLSSGSNFCSEQKTSGIINKFCSAK 340
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 341 DSEHNEKYEDTFLESEEIGTKVEVVERKEHLHTDILK------RGSEMDNNCSP 388
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           ---HHHHSESTKAPMPLLPSPPPPEPESSEDPISP 400
                                                                                                                                                                                                                                                  76 ATAGTECRKSV-----PCGWERVVKQRLFGKTAGRFDVYFISPQGLKFRSKS 122
                                                                                                                                                                                                                                                                          77 ASASPKORRSIIRDRGPMYDDPTLPEGWTRKLKORKSGRSAGKYDVYLINPOGKAFRSKV 136
                                                                                                                                                                                                                                                                                                      123 SLANYLHKNGETSLKPEDFDFTVLSKRGIKSRYKDCSMAALTSHLONQSNNSNWNLRTRS 182
                                                                                                                                                                                                                                                                                                                                                                                                                                    242 SAQVMVIK--RPGRKRKAEADPQAIPKKRGRKPGSVVAAAAAEAKKKAVKESSIRSVHET 299
                                                 PRINTS; PR00929; ATHOOK.
SMART; SM00391; MBD; 1.
Transcription regulation; Repressor; DNA-binding; Nuclear protein.
                                                                                                                                                                                                                                                                                                                                                                                  190 R-----PKGSGTGRPKAAASEGVQVKRVLEKSPGKLVVKMPFQASPGGKGEGGGATT
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                Murphy L., Richards C., Harris D., Barrell B.G., Rajandream M.A., Walsh S.V.;
Submitted (JUL-1995) to the EMBL/GenBank/DDBJ databases.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  Bauer A., Koelling R.;
"Characterization of the SAC3 gene of Saccharomyces cerevisiae.";
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       Saccharomyces cerevisiae (Baker's yeast).
Eukaryota: Fungi; Ascomycota; Saccharomycotina; Saccharomycetes;
Saccharomycetales; Saccharomycetaceae; Saccharomyces.
                                                                                                                                                                                                6.5%; Score 198.5; DB 1; Length 484; 24.1%; Pred. No. 1.1e-05;
                                                                                                                                                                                                                          50; Mismatches 154; Indels
                                                                                                                                                        62FD228F0118A49F CRC64;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     01-NOV-1995 (Rel. 32, Created)
01-OCT-1996 (Rel. 34, Last sequence update)
01-OCT-1996 (Rel. 39, Last annotation update)
LEUCINE PERMEASE TRANSCRIPTIONAL REGULATOR.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               PRT; 1301 AA
                                                                                                    POLY-ALA.
POLY-HIS.
POLY-PRO.
                                                                                                                                             POLY-THR
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            SAC3 OR LEP1 OR YDR159W OR YD8358.13.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         SEQUENCE FROM N.A. MEDLINE=97027306; Pubmed=8873450;
                                                                                                                                                        484 AA; 52307 MW;
         InterPro; IPR000637; AT_hook.
InterPro; IPR001739; MBD.
Pfam; PF01429; MBD; 1.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             401 PEPQDLSSSICKEEKMPR 418
                                                                                                                                                                                                            24.1%;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     389 TRKDFTGEKIFQEDTIPR 406
                                                                                                                                                                                                                          91; Conservative
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               STANDARD;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            Yeast 12:965-975(1996).
                                                                                                                  372
391
443
                                                                                           149
MGD; MGI:99918; Mecp2.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      STRAIN-S288C / AB972;
                                                                                                                                                                                                              Similarity
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           360 SPРККЕННН----
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        SEQUENCE FROM N.A.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                NCBI_TaxID=4932;
                                                                                                                  366
384
                                                                                           96
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               SAC3_YEAST
P46674;
                                                                                                                                             DOMAIN
SEQUENCE
                                                                                                                                                                                                 Query Match
                                                                                                                                                                                                               Best Local
                                                                                            DOMAIN
                                                                                                                                  DOMAIN
                                                                                                       DOMAIN
                                                                                                                    DOMAIN
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  YEAST
                                                                                                                                                                                                                          Matches
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  Db
 δ
                                                                                                                                                                                                                                                                                                                               qq
                                                                                                                                                                                                                                                                                                                                                        ŏ
                                                                                                                                                                                                                                                                                                                                                                                qq
                                                                                                                                                                                                                                                                                                                                                                                                           ò
                                                                                                                                                                                                                                                                                                                                                                                                                                   g
                                                                                                                                                                                                                                                                                                                                                                                                                                                             ŏ
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 δλ
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        Ω
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     ò
```

```
This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See http://www.isb-sib.ch/announce/
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        25;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         758 KDFI------DTMTRELYDAFLHERLYLIYMDSRAELKRNSTLKKKFFEKWQASYSQAK 810
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             Query Match 4.7%; Score 145; DB 1; Length 1301;
Best Local Similarity 21.9%; Pred. No. 0.11;
Matches 116; Conservative 61; Mismatches 185; Indels 168; Gaps
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     123 SLANYLHKNGETSLKPEDFDFTVLSKRGIKSRYKDC-----S 159
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       510 TLQHYSHKLSETQPLKKTY-LTCLERRLQKTTYKGLINGGEDNLASSVYVKDPKKDRIPS 568
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             160 MAALTSHLQNQSNNSNWNLRTRSKCK-----KDVFMPPSSSSELQESRGLSNFTSTHL 212
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      569 IADQSFLMENFQNNYNEKLNQNSSVKPQINTSPKRVATRPNHFPFSQESKQLPQISQSHT 628
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 213 LLKEDEGVDDVNFRKVRKPKGKVTILKGIPIKKTKKGCRKSCSGFVQSDSKRESVCNKAD 272
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   629 L-----STNPLLTPQVHGDLSEQKQQQIKTVTDG----GSPFVFDQSAQNSTVEASK 676
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         273 AESEPVAQKSQLDRTVCISDAGACGETLSVTSEENSLVKKKERSLSSGSNFCSEQKTSGI 332
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  677 A-------EEEKT--- 712
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            333 INKECSAKDSEHNEKYEDTFLESEEIGTKV--EVVERKEHLHTDILKRGSEMDNNCSPTR 390
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   713 ----QLKKKQEN---ADKQVITEQIANDLVKEVVNSSV---ISIVKREFSEAN----YR 757
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     391 KDFTGEKIFQEDTIPR----TQIERRKTSLYFSSKYNKEALSPPRRKAFKKW----- 438
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             439 -----TPPRSPFNLVQE------TLFHDPWKLLIATIFLNRTSGKMAI--PV-- 477
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              811 KNRILEEKKREEIKLVSHQLGVPGFKKSTCLFRTPYKGNVNSSFMLSSSDKNLIFSPVND 870
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     --LWKFLE-----KYPSAEVARTADWRDVSELLKPLGLYDLR 512
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           871 EFNKFATHLTKISKLWRPLEMQSIYYDNLTKKFP-----SNSLTPANLF-IY 916
                                                                                      Stella C.A., Korch C., Ramos E.H., Mattoon J.R.; "Cloning and sequencing of LEP1, a gene associated with leucine
                                                                                                                                                                                                                                                              -!- FUNCTION: POTENTIAL REGULATOR OF LEUCINE PERMEASE GENE(S)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             513 AKTIVKFSDEYLTKQWKYPIELHGIGKYGND--SYRIFCVNEWKQVHPED 560
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      Transcription regulation; Nuclear protein.
SEQUENCE 1301 AA; 149568 MW; 0679DB1673DDACEB CRC64;
                                                                                                                                                                                                                                                                                                  EXPRESSION. REQUIRED FOR NORMAL MITOSIS.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    PRT; 34.18 AA
                                                                                                                                                                                                                                                                                                                                                                                         -! - SIMILARITY: BELONGS TO THE SAC3 FAMILY.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    or send an email to license@isb-sib.ch).
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    BRC2_HUMAN STANDARD; PP51587; Q1.3879; O00183; O15008;
SEQUENCE OF 189-1301 FROM N.A.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           EMBL; Z47805; CAA87767.1; -.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        EMBL; U35227; AAA79056.1; -. SGD; S0002566; SAC3.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   EMBL; Z50046; CAA90379.1;
                                                                                                                                                                                                                     Yeast 11:460-460(1995).
                                             STRAIN-S288C / DBY939;
                                                                                                                                                                            transport.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           BRC 2_HUMAN
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       RESULT
```

```
MEDLINE=96112016; PubMed=8524414;
MOOSter R. Bignell G. Lancaster J., Swift S., Seal S., Manqion J.,
Collins N., Gregory S., Gumbs C., Micklem G., Barfoot R., Hamoudi R.,
Patel S., Rice C., Biggs P., Hashim Y., Smith A., Connor F.,
Arason A., Gudmundsson J., Ficenec D., Kelsell D., Ford D., Tonin P.,
Bishop D.T., Spurr N.K., Ponder B.A.J., Eeles R., Peto J., Devilee P.,
Cornelisse C., Lynch H., Narcod S., Lenoir G., Eglisson V.,
Barkadottir R.B., Easton D.F., Bentley D.R., Futreal P.A.,
Ashworth A., Stratton M.R.;
"Identification of the breast cancer susceptibility gene BRCA2.";
Nature 378:789-792(1995).
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               MEDLINE=96172838; PubMed=8589730;

Tavtigian S.V., Simard J., Ronmens J., Couch F., Shattuck-Eidens D., Tavtigian S.V., Simard J., Ronmens J., Couch F., Stoppa-Lyonnet D., Neuhausen S., Merajver S., Thorlanderius S., Offit K., Stoppa-Lyonnet D., Belanger C., Bell R., Berry S., Bogden R., Chen O., Davis T., Dumont M., Frye C., Hattier T., Jammulapati S., Janecki T., Jiang P., Kehrer R., Leblanc J.-F., Mitchell J.T., McArthur-Morrison J., Ruyyen K., Peng Y., Samson C., Schroeder M., Snyder S.C., Steele L., Stringfellow M., Stroup C., Swedlund B., Swensen J., Teng D., Thomas A., Tran T., Tranchant M., Weaver-Feldhaus J., Mong A.K.C., Shizuya H., Eyfjord J.E., Cannon-Albright L., Labrie F., Skolnick M.H., Weber B., Kamb A., Goldar D.E., "The complete BRCA2 gene and mutations in chromosome 13g-linked
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      Couch F.J., Farid L.M., Deshano M.L., Tavtigian S.V., Calzone K., Campeau L., Peng Y., Bogden B., Chen Q., Neuhausen S., Shattuck-Fildens D., Godwin A.Y., Daly M., Radford D.M., Sedlacek S., Rommens J., Simard J., Garber J., Merajver S., Weber B.L.; BRCA2 germline mutations in male breast cancer cases and breast
                                                                                                                                            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 Lancaster J.W., Wooster R., Mangion J., Phelan C.M., Cochran C., Gumbs C., Seal S., Barfoot R., Collins N., Bignell G., Patel S., Hamoudi R., Larsson C., Wiseman R.W., Berchuck A., Iglehart J.D., Marks J.R., Ashworth A., Stratton M.R., Futreal P.A.; "BRCAZ mutations in primary breast and ovarian cancers."; Nat. Genet. 13:238-240(1996).
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   VARIANT OC HIS-2787, AND VARIANTS ASN-372; MET-1915 AND VAL-2466
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          MEDLINE-96275740; PubMed-8665505;
Takahashi H., Chiu H.-C., Bandera C.A., Behbakht K., Liu P.C.,
Couch F.J., Weber B.L., LivJolsi V.A., Furusato M., Rebane B.A.,
Cardonick A., Benjamin I., Morgan M.A., King S.A., Mikuta J.J.,
Rubin S.C., Boyd J.;
"Mutations of the BRA2 gene in ovarian carcinomas.";
Cancer Res. 56:2738-2741(1996).
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          VARIANTS ASN-372; ASP-991; SER-1147; MET-1915 AND CYS-2034. MEDLINE=96241588; PubMed=8673091;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           Hunt S., Mcmurray A., Williamson H.;
Submitted (MAY-1997) to the EMBL/GenBank/DDBJ databases.
                    -2001 (Rel. 40, Last sequence update)
-2001 (Rel. 40, Last annotation update)
CANCER TYPE 2 SUSCEPTIBILITY PROTEIN.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        VARIANT GLU-3095.
MEDLINE=96225456; PubMed=8640235;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       MEDLINE=96225457; PubMed=8640236;
(Rel. 34, Created)
(Rel. 40, Last seq
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     Nat. Genet. 13:123-125(1996).
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      Nat. Genet. 12:333-337(1996).
                                                                                                                     Homo sapiens (Human).
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    SEQUENCE FROM N.A.
                                                                                                                                                                                                                                              SEQUENCE FROM N.A.
                                                                                                                                                                                             NCBI_TaxID=9606;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        families.
                    20-AUG-2001
20-AUG-2001
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               kindreds
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  VARIANTS
                                                                       BREAST
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      cancer
DDDTT
BDDTT
BDTT
BDDTT
BDDT
BDDTT
BDTT
BDDTT
BDTT
```

```
WARDINES BC L-32; R-53; L-81; R-201; A-211; S-222 AND T-3118.

MEDLINE=98272917; PubMed=9609997;
Katagiri T., Kasumi F., Yoshimoto M., Nomizu T., Asaishi K., Abe R., Tsuchiya A., Sugano M., Takai S., Yoneda M., Fukutomi T., Nanba K., Makita M., Okazaki H., Hirata K., Okazaki M., Prutusuma Y., Morishita W., Okazaki M., Yamazaki M., Prutusuma Y., Houga S., Shimizu T., Toda M., Yamazaki Y., Uchida T., Kajiwara T., Houga S., Shimizu T., Toda M., Yamazaki Y., Uchida T., Kunitomo K., Sonoo H., Kurebayashi J.-I., Shimotsuma K., Nakamura Y., Miki Y.; Hidip proportion of missense mutations of the BRCAl and BRCA2 genes in Japanese breast cancer families."
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            MEDILNE=98316775; PubMed-9654203; Ganguly A.; Ganguly T., Dhulipala R., Godmilow L., Ganguly A.; Ganguly T., Dhulipala R., Godmilow L., Ganguly A.; "High throughput fluorescence-based conformation sensitive gel electrophoresis (F-CSGE) identifies six unique BRCA2 mutations and an overall low incidence of BRCA2 mutations in high-risk BRCA1-negative breast cancer families.";
Teng D.H.-F., Bogden R., Mitchell J., Baumgard M., Bell R., Berry S., Davis T., Ha P.C., Kehrer R., Jammulapati S., Chen Q., Offit K., Skolnick M.H., Tavtigian S.V., Jhanwar S., Swedlund B., Wong A.K.C.,
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           MEDLINE=99254821; PubMed=10323242;
Li S.S.-L., Tseng H.-M., Yang T.-P., Liu C.-H., Teng S.-J.,
Huang H.-W., Chen L.-W., Kao H.-W., Chen J.H., Tseng J.-N., Chen A.,
Hou M.-F., Huang T.-J., Chang H.-T., Mok K.-T., Tsai J.-H.;
"Molecular characterization of germline mutations in the BRCA1 and
BRCA2 genes from breast cancer families in Taiwan."; the BRCA1 and
Hum. Genet. 104:201-204(1999).
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              MEDLINE×20455732; PubMed*10486320;
Gayther S.A., Russell P., Harrington P., Antoniou A.C., Easton D.F.,
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         MEDLINE 99136688; pubMed-9971877; Wapter T.M.U., Hittenlehner K., Shen P., Moeslinger R., Muhr D., Fleischmann E., Concin H., Doeller W., Haid A., Lang A.H., Mayer P., Fleischmann E., Concin H., Doeller W., Haid A., Lang A.H., Mayer P., Underhill P., Mountain J., Stierer M., Zielinski C., Oefner P., "Global sequence diversity of BRCA2: analysis of 71 breast cancer families and 95 control individuals of worldwide populations."; Hum. Mol. Genet. 8:413-423(1999).

-:-FUNCTION: MAY PARTICIPATE IN A PATHWAY ASSOCIATED WITH THE ACTIVATION OF DOUBLE-STRAND BREAK REPAIR AND/OR HOMOLOGOUS RECOMBINATION.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   "The contribution of germline BRCA1 and BRCA2 mutations to familial ovarian cancer: no evidence for other ovarian cancer-susceptibility
                                                                                                                                                                                                                                                                                                                                                                                                                                                                         Miki Y., Katagiri T., Kasumi F., Yoshimoto T., Nakamura Y.; "Mutation analysis in the BRCA2 gene in primary breast cancers.";
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        Vehmanen P., Friedman L.S., Berold H., Sarantaus L., Pyrhoenen S. Ponder B.A.J., Muhonen T., Nevanlinna H.; "A low proportion of BRCA2 mutations in Finnish breast cancer families.";
                                                                                                                                                                                               "Low incidence of BRCA2 mutations in breast carcinoma and other
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        VARIANT BC ASP-2089, AND VARIANT VAL-3412.
MEDLINE=97294396; PubMed=9150152;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             VARIANTS OC PRO-75; HIS-2502 AND HIS-3098
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              VARIANT BC AND PANCREAS CANCER TRP-554.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            Am. J. Hum. Genet. 60:1050-1058(1997).
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       J. Hum. Genet. 65:1021-1029(1999).
                                                                                                                                                                                                                                                                                                                                                                                                                                      MEDLINE=96225458; PubMed=8640237;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 breast cancer families.";
Hum. Genet. 102:549-556(1998).
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    Hum. Genet. 43:42-48(1998).
                                                                                                                                                                                                                                                                                        Nat. Genet. 13:241-244(1996).
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           Nat. Genet. 13:245-247(1996).
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      VARIANTS BC, AND VARIANTS
                                                                                                                                                                                                                                                                                                                                                                                VARIANT ASN-2415.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           Ponder B.A.J.;
                                                                                                                                                                                                                                                cancers."
                                                                                                                                              Kamb A.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               genes.
```

```
This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation the European Bioinformatics Institute. There are no restrictions on its
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   23;
                                                                                                                                                                                         use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See http://www.isb-sib.ch/announce/
          -i- TISSUE SPECIFICITY: HIGHEST LEVELS OF EXPRESSION IN BREAST AND THYMUS, WITH SLIGHTLY LOWER LEVELS IN LUNG, OVARY, AND SPLEEN.
-i- DISASE: BREAST CANCER (BC) IS AN EXTREMELY COMMON MALLIGANOX, AFFECTING ONE IN EIGHT WOMEN DURING THEIR LIFETIME. A POSITIVE FAMILY HISTORY HAS BEEN IDENTIFIED AS MAJOR CONTRIBUTOR TO RISK OF DEVELOPMENT OF THE DISEASE, AND THIS LINK IS STRKING FOR EARLY-ONSET BREAST CANCER. MUTATIONS IN BRCA2 ARE THOUGHT TO BE RESPONSIBLE FOR SOME INHERITED BREAST CANCER. IT IS LINKED WITH MALE BREAST CANCER.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     972 KSDISLNIDKIPEKNNDYMNKWAG----LLGPISNHSFGGSFRTASNKEIKLSEHNIKK 1026
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         1027 SKMFFK-----DIEEQYPTSL---ACVEIVNTLALDNQKKLSKPQSIN-TVSAHLQSSVV 1077
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               1078 VSDCKNSHITPQMLFSKQDFNSNHNL-------TPSQKAEITE-----LSTIL 1118
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      1119 ----EESGSQFEFTQFRKPSYILQKSTFEVPENQMTILKTTSEECRDADLHVIMNAPSIG 1174
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      155 YKDCSMAALTSHL--QNQSNNSNWNLRTRSKCKKDVFMPPSSSSELQESRGLSNFTSTHL 212
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            213 LLKEDEGVDDVNFRKVRK------289
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 96 QRLFGKTAGRFDVYFISPQGLKFRSKSSLANYLHKNGETSL-KPEDFDFTVLSKRGIKSR 154
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   Gaps
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             36 KEDVAMELERVGEDEEQMMIKRSSECNPLLQEPIASAQFGATAGTECRKSVPCGWERVVK 95
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 67; Mismatches 158; Indels 138;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       4.7%; Score 145; DB 1; Length 3418;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   Pred. No. 0.39;
                                                                                                                                                                                                                                 or send an email to license@isb-sib.ch).
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                interPro; IPR002093; BRCA2_repeat.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         PROSITE; PS50138; BRCA2_REPEAT; 8.
 SUBUNIT: INTERACTS WITH RAD51.
                                                                                                                                                                                                                                                                                                                                                                                                                                                          JOINED.
                                                                                                                                                                                                                                                                                        JOINED.
                                                                                                                                                                                                                                                                                                                                                           JOINED.
JOINED.
                                                                                                                                                                                                                                                                                                                                                                                                                               JOINED.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             JOINED.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      JOINED.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    JOINED.
                                                                                                                                                                                                                                                                             JOINED.
                                                                                                                                                                                                                                                                                                                     JOINED.
                                                                                                                                                                                                                                                                                                                                  JOINED.
                                                                                                                                                                                                                                                                                                                                               JOINED.
                                                                                                                                                                                                                                                                                                                                                                                        JOINED.
                                                                                                                                                                                                                                                                                                                                                                                                   JOINED.
                                                                                                                                                                                                                                                                                                                                                                                                                  JOINED.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 JOINED.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 JOINED.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          JOINED.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            Pfam; PF00634; BRCA2_repeat;
                                                                                                                                                                                                                                                             EMBL; X95152; CAA64484.1; -
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 20.98;
                                                                                                                                                                                                                                                                              CAA64484.1;
                                                                                                                                                                                                                                                                                        CAA64484.1;
CAA64484.1;
                                                                                                                                                                                                                                                                                                                                                            CAA64484.1;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     CAA64484.1;
CAA64484.1;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 CAA64484.1;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      273359; CAA97728.1;
                                                                                                                                                                                                                                                                                                                     CAA64484.1;
                                                                                                                                                                                                                                                                                                                               CAA64484.1;
                                                                                                                                                                                                                                                                                                                                               CAA64484.1;
                                                                                                                                                                                                                                                                                                                                                                                                     CAA64484.1;
                                                                                                                                                                                                                                                                                                                                                                                                                                             CAA64484.1;
                                                                                                                                                                                                                                                                                                                                                                                                                                                          CAA64484.1:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      CAA64484.1;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   CAA64484.1;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 CAA64484.1;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          CAA64484.1;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          CAA64484.1;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            CAA64484.1;
AAB07223.1;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      CAA98995.1;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                       CAA64484.1;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   Best Local Similarity 20.9
Matches 96; Conservative
                                                                                                                                                                                                                                                                                                                                                                                                                             CAA64484.1
                                                                                                                                                                                                                                                                                                                                                                           CAA64484.1
                                                                                                                                                                                                                                                                                                                  X95156;
X95157;
X95158;
                                                                                                                                                                                                                                                                                                                                                           X95159;
X95160;
                                                                                                                                                                                                                                                                                          X95154;
                                                                                                                                                                                                                                                                                                                                                                                                  x95162;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               x95176;
                                                                                                                                                                                                                                                                                                     x95155;
                                                                                                                                                                                                                                                                                                                                                                                      x95161;
                                                                                                                                                                                                                                                                                                                                                                                                                               X95164;
                                                                                                                                                                                                                                                                                                                                                                                                                                             x95165;
                                                                                                                                                                                                                                                                                                                                                                                                                                                          x95166;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      x95168;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   x95169;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              x95171;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           043746;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                         x95167;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        274739;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       Query Match
                                                                                                                                                                                                                                                                                                                                            EMBL;
                                                                                                                                                                                                                                                                                                                                                                        EMBL;
                                                                                                                                                                                                                                                                                          EMBL;
                                                                                                                                                                                                                                                                                                       EMBL;
                                                                                                                                                                                                                                                                                                                     EMBL;
                                                                                                                                                                                                                                                                                                                                 EMBL;
                                                                                                                                                                                                                                                                                                                                                            EMBL;
                                                                                                                                                                                                                                                                                                                                                                                                                                             EMBL;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                       EMBL;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 EMBL;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  EMBL;
                                                                                                                                                                                                                                                                                                                                                                                        EMBL;
                                                                                                                                                                                                                                                                                                                                                                                                     EMBL;
                                                                                                                                                                                                                                                                                                                                                                                                                   EMBL;
                                                                                                                                                                                                                                                                                                                                                                                                                               EMBL;
                                                                                                                                                                                                                                                                                                                                                                                                                                                          EMBL;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      EMBL;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 EMBL;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              EMBL;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          EMBL;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          EMBL;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     EMBL;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 EMBL;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              EMBL;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            EMBL;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         EMBL;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        EMBL;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    qq
g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         q
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      δλ
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            ó
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           δ
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 δ
```

```
between the Swiss Institute of Bioinformatics and the EMBL outstation—
the European Bioinformatics Institute. There are no restrictions on its
use by non-profit institutions as long as its content is in no way
modified and this statement is not removed. Usage by and for commercial
entities requires a license agreement (See http://www.isb-sib.ch/announce/
or send an email to license@isb-sib.ch).
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              This SWISS-PROT entry is copyright. It is produced through a collaboration
                                1175 QVDSSKQFEGTVEIKRKFAGLLKNDCNKSASGYLTDE-----NEVGFRGFYSAHGTKL 1227
                                                                                                        1228 NVSTEALQKAVKLFSDIENISEETS---AEVHPISLSSSKCHD----SVVSMF---KIEN 1277
                                                                                                                                                                            1278 HNDK---TVSEKNNKCQLILQNNIEMTTGTFVEEI-----TENYKRNTENEDNKYTAA 1327
                                                                    285 D-RTVCISDAGACGETLSVTSEENSLVKKKERSLSSGSNFCSEQKTSGIINKFCSAKDSE 343
                                                                                                                                            344 HNEKYEDTFLESEE------IGTKVEVVERKEHLHTDILKRGSEMDNN--CSP 388
240 -----GIPIKK----TKKGCRKSCSGFVQSDSKRESVCNKADAESEPVAQKSQL 284
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       "gar2 is a nucleolar protein from Schizosaccharomyces pombe required for 18S rRNA and 40S ribosomal subunit accumulation."; Nucleic Acids Res. 23:1912-1918(1995).
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               Brown D., Churcher C.M., Barrell B.G., Rajandream M.A., Wood V.; Submitted (SEP-1997) to the EMBL/GenBank/DDBJ databases.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   Gulli M.-P., Girard J.-P., Zabetakis D., Lapeyre B., Melese T., Caizergues-Ferrer M.;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     Ribosome biogenesis; RNA-binding; Nuclear protein; Repeat;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         -!- SIMILARITY: BELONGS TO THE GAR FAMILY.
-!- SIMILARITY: CONTAINS 2 RNA RECOGNITION MOTIFS (RRM).
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          Eukaryota; Fungl; Ascomycota; Schizosaccharomycetes;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       Schizosaccharomycetales; Schizosaccharomycetaceae;
                                                                                                                                                                                                                                                 1328 SRNSHNLEFDGSDSSKNDTV---CIHKDETDLLFTDQHN 1363
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          RNA-BINDING (RRM) 1. RNA-BINDING (RRM) 2.
                                                                                                                                                                                                                389 TRK----DFTGEKIFQEDTIPRTQIERRKTSLYFSSKYN 423
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              S -> P (IN REF. 1).
                                                                                                                                                                                                                                                                                                                                              GAR2_Scurc
P41891; 013707;
01-N0V-1995 (Rel. 32, Created)
30-MAY-2000 (Rel. 39, Last sequence update)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         -!- SUBCELLULAR LOCATION: NUCLEAR; NUCLEOLUS.
                                                                                                                                                                                                                                                                                                                                            500 AA
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             GLU/SER-RICH.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                       Schizosaccharomyces pombe (Fission yeast).
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    MEDLINE=95319932; PubMed=7596817;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     PROSITE; PS50102; RRM; 2.
PROSITE; PS00030; RRM_RNP_1; 2.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             EMBL; AL163191; CAB86413.1;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              InterPro, IPR000504, RRM.
Pfam; PF00076, rrm; 2.
SMART; SM00360; RRM; 2.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          EMBL; 248166; CAA88179.1;
                                                                                                                                                                                                                                                                                                                                            STANDARD;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        CONTAINING 18S RRNA
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          341
443
339
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            Schizosaccharomyces.
                                                                                                                                                                                                                                                                                                                                                                                                                                                      GAR2 OR SPAC140.02.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              HSSP; P09651; 1HA1.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               SEQUENCE FROM N.A.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 SEQUENCE FROM N.A.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          263
366
339
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             NCBI_TaxID=4896;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            rRNA processing.
                                                                                                                                                                                                                                                                                                                                          GAR2_SCHPO
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  STRAIN=972;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    STRAIN=972
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                CONFLICT
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            DOMAIN
DOMAIN
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             DOMAIN
                                                                                                                                                                                                                                                                                                                       GAR2_SCHPO
                                Op
                                                                                                      qq
                                                                                                                                                                        Q
                                                                                                                                                                                                                δλ
                                                                                                                                                                                                                                                                                                                                                                              QΥ
                                                                                                                                          ŏ
```

```
SMART; SM00249; PHD; 1. SMART; SM00317; SET; 1.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         NCB1_Tax1D=9606
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            387 Q----
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               CYL1_HUMAN
P35663;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               532 I 532
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          L 478
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     RESULT 8
CYL1_HUMAN
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           478
                                                                                                                             Matches
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  DR
DR
SQ
                                                                                                                                                         ŏ
                                                                                                                                                                                  qq
                                                                                                                                                                                                               δy
                                                                                                                                                                                                                                         qq
                                                                                                                                                                                                                                                                     ŏ
                                                                                                                                                                                                                                                                                                q
                                                                                                                                                                                                                                                                                                                            Οý
                                                                                                                                                                                                                                                                                                                                                       qq
                                                                                                                                                                                                                                                                                                                                                                                    ó
                                                                                                                                                                                                                                                                                                                                                                                                               QQ
                                                                                                                                                                                                                                                                                                                                                                                                                                           ó
                                                                                                                                                                                                                                                                                                                                                                                                                                                                    QQ
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            qq
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        Qγ
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     Ωp
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 ó
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation the Buropean Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (see http://www.isb-sib.ch/announce/or send an email to license@isb-sib.ch).
                                                                       11;
                                                                                                                                                                                    361
                                                                                                                                                                                                                                           181
                                                                                                                                                                                                                                                                       420
                                                                                                                                                                                                                                                                                                                              421 KYNKEALSPPRRKAFKKWTPPRSPFNLVQETLFHDPWKLLIATIFLNRTSGKMAIPVLWK 480
                                                                                                                                                                                                                                                                                                                                                        281
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 37.5 kb region of yeast chromosome X includes the SME1, MEF2, GSH1
                                                                                                 185 KKDVFMPPSSSSELQESRGLSNFTSTHLLLKEDEGVDDVNFRKVRKPKGKVTILKGIPIK 244
                                                                                                                                                        245 KTKKGCRKSCSGFVQSDSKRESVCNKADAESEPVAQKSQLDRTVCISDAGACGETLSVTS 304
                                                                        Gaps
                                                                                                                             19 KKGAIEKPSKSKKITKEAAKE----IAKQSSKTDVSPKKSKKEAKRASSPE--PSK 68
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             and CSD3 genes, a TCP-1-related gene, an open reading frame similar to the DAL80 gene, and a tRNA(Arg)."; Yeast 11:873-883(1995).
                                                                                                                                                                                                                EENSLVK----KKERSLSSGSNFCSEQKTSGIINKFCSAKDSEHNEKYEDTFLESEEIGTK
                                                                                                                                                                                                                                           123 EEEVIVKTEEKKESSSESSSESSESEEEEAVV-KIEEKKESSSDSSSESSSSESESSS
                                                                                                                                                                                                                                                                       362 VEVVERKEHLH-TDILKRGSEMDNNCSPTRKDFTGEKIFQEDTIPRTQIERRKTSLYFSS
                                                                                                                                                                                                                                                                                                                                              182 SESEEEEEVVEKTEEKKEGSSESSSDSESSSDSSSE---SGDSDSSSDSESESS---SE
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   Saccharomyces cerevisiae (Baker's yeast).
Eukaryota; Fungi; Ascomycota; Saccharomycotina; Saccharomycetes;
Saccharomycetales; Saccharomycetaceae; Saccharomyces.
                                                                     41;
                                            Length 500;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      01-NOV-1995 (Rel. 32, Last sequence update)
20-AUG-2001 (Rel. 40, Last annotation update)
HYPOTHETICAL 63.9 KDA PROTEIN IN IME2-MEF2 INTERGENIC REGION
XJL105W OR J0819.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      -i-SIMILARITY: TO YEAST YKR029C AND S.POMBE SPAC22E12.11C.
-i-SIMILARITY: CONTAINS 1 PHD-TYPE ZINC FINGER.
-i-SIMILARITY: CONTAINS 1 SET DOMAIN.
                                                                        Indels
  52987 MW; 9D37FAD0C5161A0B CRC64;
                                       / Match 4.5%; Score 138.5; DB 1; Local Similarity 23.7%; Pred. No. 0.08; les 74; Conservative 50; Mismatches 147;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      560 AA.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      PRT;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       STRAIN=S288C / FY1679;
MEDLINE=96090136; PubMed=7483851;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 01-NOV-1995 (Rel. 32, Created)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            EMBL; X85021; CAA59389.1; -. EMBL; Z49380; CAA89400.1; -.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        SGD; S0003641; YJL105W.
InterPro; IPR001965; PHD.
InterPro; IPR001214; SET.
Pfam; PF00628; PHD; 1.
Pfam; PF00856; SET; 1.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      STANDARD;
                                                                                                                                                                                                                                                                                                                                                                                      481 FLEKYPSAEVAR 492
                                                                                                                                                                                                                                                                                                                                                                                                                282 EFEEYGTIVGAR 293
  500 AA;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            SEQUENCE FROM N.A.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               NCBI_TaxID=4932;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     Rasmussen S.W.;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      YJK5_YEAST
  SEQUENCE
                                               Query Match
                                                                                                                                                                                                                                                                                                                                                                                                                                                                         YJK5_YEAST
                                                                                                                                                                                                                305
                                                                        Matches
    So
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   g
                                                                                                                                                           ò
                                                                                                                                                                                     q
                                                                                                                                                                                                                  ò
                                                                                                                                                                                                                                         g
                                                                                                                                                                                                                                                                       ð
                                                                                                                                                                                                                                                                                                   a
                                                                                                                                                                                                                                                                                                                               ò
                                                                                                                                                                                                                                                                                                                                                                                      ŏ
                                                                                                                                                                                                                                                                                                                                                                                                                8
```

```
24;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 Hess H., Held H., Franke W.W.;
"Molecular characterization of mammalian cylicin, a basic protein of the sperm head cytoskeleton.";
"Incell Biol. 122:1043-1052(1993).
-!- FUNCTION: POSSIBLE ARCHITECTURAL ROLE DURING SPERMATOGENESIS. MAY BE INVOLUDE IN SPERMATID DIFFERENTIATION.
-!- SUBCELLULAR LOGATION: CALYX: SPERM HEAD CYTOSKELETAL STRUCTURE.
-!- TISSUE SPECIFICITY: TESTIS.
                                                                                                                                                                                                                                                                                                                118 FRSKSSLANYLHKNGETSLKPEDFDFTVLSKRGIKSRYKDCSMAALTSHL--QNQSNNSN 175
                                                                                                                                                                                                                                                                                                                                                                                                                                                        176 WNLRTRSKCKKDVFMPPSSSSELQESRGLSNFTSTHLLLKEDEGVDDVNFRKVRKPKGKV 235
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    ---FIQCNKCKTWQHKLCYAFKKSDPIKRDFVCKRCDSDTK--VQVNQVKPMIFPRKMG 228
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    ----TLSVTSEENSLVKK-KERSLS 318
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 229 DERLFOFSSIVTTSASNTNQHQQSVNNIEEQPKKRQLHYTAPTTENSNSIRKKLRQEKLV 288
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     319 SGSNFCSEQKTSGIINKFCSAKD-----SEHNEKYEDTFLESEEIGTKVEVVERKEH 370
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             289 VSSHFLKP-----LLNEVSSSNDTEFKAITISEYKDKYVKMFIDN------H 329
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          371 LHTDIL------KRGSEMDNNCSPTRKDFTGEKIFQEDTIPRTQ-IERRKTSLYFSSKY 422
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        423 NKEALSPPRRKAFKKWTPPRSPFNLVQET----LFHDPWKLLIATIFLNRTSGKMAIPVL 478
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               ------TDPNNDYRLMGTTKPKVLFHPHWPLYIDS---RETGG-----L 422
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      479 WKFLEK--YPSAEVARTADWRDVSELLKPLGLYDLRAKTIVKF-----SDEYLTKQWKYP 531
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          82 FONDRGIFNHHSSSGSSKT----ASTNKRGIAAAVALATAATIPFPLKKONODDNSK 134
                                                                                                                                                                                                                                               Indels 150; Gaps
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     135 VSV-THNESSKENKITPSMRAE-----DNKPKNGCICGSSDSKDEL------
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            236 TILKGIPIKKTKKGCRKSCSGFVQSDS-KRESVCNKADAESEPVAQKSQLD-----
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                Eukaryota, Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
                                                                                                                                                                        4.5%; Score 137; DB 1; Length 560; 18.7%; Pred. No. 0.12;
PROSITE: PS50280; SET; 1.
Hypothetical protein; Zinc-finger.
SEQUENCE 560 AA; 63855 MW; 9A7BAECE97318AFA CRC64;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    01-00N-1994 (Rel. 29, Last sequence update)
20-AUG-2001 (Rel. 40, Last annotation update)
CYLICIN I (MULTIPLE-BAND POLYPEPTIDE I) (FRAGMENT).
                                                                                                                                                                                                                                               Mismatches 161;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       598 AA
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           PRT;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               TISSUE=Testis;
MEDLINE=93359502; PubMed=8354692;
                                                                                                                                                                                                                                                   80;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           01-JUN-1994 (Rel. 29, Created)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    286 -----RTVCISDAGACGE----
                                                                                                                                                                                                                                                   90; Conservative
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           STANDARD;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            CYLC1 OR CYL1 OR CYL.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   Homo sapiens (Human)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        SEQUENCE FROM N.A.
                                                                                                                                                                                   Query Match
Best Local Similarity
```

```
NP_BIND
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             DOMAIN
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              DOMAIN
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    EMBL;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          Matches
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      132
            DDRR W B DDR R W
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    ŏ
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     QO
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      ò
                                         This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation - the Buropean Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See http://www.isb-sib.ch/announce/or send an email to license@isb-sib.ch).
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            21;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         118 FRSKSSLANYLHKNGETSLKPEDFDFTVLSKRGIKSRYKDCSMAALTSHLQN--QSNNSN 175
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           70 GGTPLKKDSKKKGGSYATNPESKQIVEEKTKRQNEAD-KTPLKSSHENEQSKKSKSSSET 128
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    172 KRSKTSNDPISEICSENSL---NVDFLMLVGQ-----SDDESINFDAWLRNYSQNNSKN 222
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         223 YSLKYTKYTKKDTKKNAKKSSDAESEDSKDAKKDSKKVKKNVKKDDKKKDVKKDTESTDA 282
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      319 SGSNFCSEQKTSGIINKFCSAKDSEHNEKYEDT--FLESEEIGTKVEVVERKEHLHTDIL 376
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               62 NPLLQEPIASAQFGATAGTECRKSVPCGWERVVKQRLFGKTAGRFDVYFI----SPQGLK 117
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             129 NP-----ESQNSKTVSKNCSQ------KDKKDSKNSKKTNTEFLHTKNNPKKDL 171
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       224 -----NFRKVRKPKGKVTILKGIPIKKTKK-------GCRKSCSGFVQS 260
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      261 DSKRES---VCNKADAESEPVAQKSQLDRTVCISDAGACGETLSVTSEENSLVKKKERSLS 318
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   393 ADSEPKGDSK-KGKKDEKKGKKDSKKDDKKKDAKKNAESTEMESDLELKKDKKHSKE--- 448
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       DKKKDAKKITFSTDSESELESKESOKDEKKDKKD-----SKTDNKKS-VKNDEESTD 392
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            Gaps
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                2 GTTGLESLSLGDRGAAPTVTSSERLVPDPPNDLRKEDVAMELERVGEDEEQMMIKRSSEC 61
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  377 KRGSEMDNNCSPTRKDFTGEKIFQEDTIPRTQIERRKTSLYFSSKYNKEALSPPRRK 433
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   449 KKGSKKDIK-KDARKD-TESTDAEFDESSKTGF-KTSTKIKGSDTESEESLYKPGAK 502
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        176 WNLRTRSKCKKDV--FMPPSSSSELQESR-GLSNFTSTHLLLKEDEGVDDV-----
                                                                                                                                                                                                                                                                                                                                           Cytoskeleton; Structural protein; Repeat; Sperm; Spermatogenesis.
NON_TER
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      89;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        DB 1; Length 598;
-!- DEVELOPMENTAL STAGE: SPECIFIC TO LATE SPERMATOGENESIS.
                                                                                                                                                                                                                                                                                                                                                                                               APPROXIMATE TANDEM REPEATS.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 72; Mismatches 205; Indels
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              062BA2E2D2AB61F7 CRC64;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     01-FEB-1996 (Rel. 33, Created)
01-OCT-1996 (Rel. 34, Last sequence update)
30 MAY-2000 (Rel. 39, Last annotation update)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                4.5%; Score 136.5; Di 23.3%; Pred. No. 0.14;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              PRT; 1447 AA
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         PRO-RICH.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              68034 MW;
                                                                                                                                                                                                                                               EMBL; 222780; CAA80457.1; -.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          Best Local Similarity 23.34
Matches 111; Conservative
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              STANDARD;
                                                                                                                                                                                                                                                                      PIR; S35920; S35920.
PIR; B40713; B40713.
MIM; 603121; -.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           598 AA;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           BUD4_YEAST
P47136;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           SEQUENCE
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   Query Match
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  REPEAT
REPEAT
                                                                                                                                                                                                                                                                                                                                                                                               DOMAIN
                                                                                                                                                                                                                                                                                                                                                                                                                        REPEAT
                                                                                                                                                                                                                                                                                                                                                                                                                                               REPEAT
                                                                                                                                                                                                                                                                                                                                                                                                                                                                    REPEAT
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                REPEAT
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             REPEAT
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      REPEAT
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       342
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  BUD4_YEAST
q
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  ŏ
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               δ
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           δ
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    Db
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           DY PY
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   Ω
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         δ
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               qq
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     δ
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             QQ
```

```
This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation-the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See http://www.isb-sib.ch/announce/or send an email to license@isb-sib.ch).
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              PTASVEDLNISTNLPGADSSQNNPVTTDADALIENDVV ->
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                SEQUENCE FROM N.A.
MEDLINE=9659555; PubMed=8707826;
MEDLINE=9695555; PubMed=8707826;
MEDLINE=9695555; PubMed=8707826;
The BUD4 protein of yeast, required for axial budding, is localized to the mother/BUD neck in a cell cycle-dependent manner.";
J. Cell Biol. 134:413-427(1996).
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        74 MSTISNESLNLGLLRVNSELEESPAAVHQERIKNSVANGALGHANSPKVLNNLK -- NMAQ 131
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 DIDKLARDEEKPVKLSSSPLKFTLKSTQPLLSYPESPIH------RSSIE-----I 176
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       95 KQRLFGKTAGRFDVYFI---SPQGLKFRSKSSLANYLHKNG---ETSLKPEDFDFTVLSK 148
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      149 RGIKSRYKI)------CSMAA-----LTSHLQNQSNNSNWNLRTRSKCKK 186
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            237 TSVDSTGRELDTKTIPELPFCMSSTPEMTPVDEKCNLPSKLLNTSNNSHSDSRSPTASVE 296
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                Local Similarity 19.5%; Pred. No. 0.73;
les 135; Conservative 105; Mismatches 269; Indels 184; Gaps
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  187 DV----FMPPSSSSELQESRGLSN--FTSTHLLLKED------EGVDDVNFRKVRKP 231
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         ANSLCGGFKHFNESSRVLIPAKIIQSLLMRMRLLKTMLC
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         1 MGTTGLESLSLG-----DRGAAPTVTSSERL------VPDPPNDLRKEDVAM 41
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        42 ELERVGEDEEQMMIKRSS-----ECNPLLQEPIASAQFGATAGTECRKSVPCGWERVV 94
                                                                                         Eukaryota; Fungi; Ascomycota; Saccharomycotina; Saccharomycetes;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 DB 1; Length 1447;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            340 340 D -> E (IN REF. 2).
1447 AA; 164485 MW; 1D056B9FF1B7067B CRC64;
                                                                                                                                                                                                                                                                                                                                                                                                                                Ramezani Rad M., Kirchrath L., Hollenberg C.P.;
Submitted (SEP-1995) to the EMBL/GenBank/DDBJ databases.
                                                                                                                 Saccharomycetales; Saccharomycetaceae; Saccharomyces.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             ASP/GLU-RICH (ACIDIC).
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     ATP (POTENTIAL)
                                                        Saccharomyces cerevisiae (Baker's yeast).
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           4.4%; Score 133; 19.5%; Pred. No. 0
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    (IN REF.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         Z49592; CAA89620.1; ALT_FRAME.
BUD4 OR YJR052W OR J1905.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   SMART; SM00233; PH; 1.
PROSITE; PS50003; PH_DOMAIN; 1.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           EMBL; U41641; AAB17116.1; -.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          EMBL; 249591; CAA89619.1; -.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          InterPro, IPR001849; PH.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       189
1182
1413
328
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           Pfam; PF00169; PH;
                                                                                                                                                                                                                                                                                                                                                                                                     SEQUENCE FROM N.A.
                                                                                                                                            NCBI_TaxID=4932;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  1175
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             1302
291
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       181
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  SGD; S0003852;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            ATP-binding.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           CONFLICT
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          SEQUENCE
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              Query Match
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            CONFLICT
```

```
SEQUENCE
                                                                                                                                                                                                                                                                        Query Match
                                                                                                                                                                                                  DOMAIN
DOMAIN
                                                                                                                                                                                      DOMAIN
                                                                                                                                                                     REPEAT
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     ANK2_HUMAN
 qq
                                                                                                                                                                                                                                                                                                                                                                                                                   qq
                                                                                                                                                                                                                                                                                                                                                                                                                                                                          q
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   qq
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  AC AC
                                                                                                                                                                                                                                                                                                                                 δ
                                                                                                                                                                                                                                                                                                                                                                                        δ
                                                                                                                                                                                                                                                                                                                                                                                                                                               ô
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            qq
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  q
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               Ö
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            qq
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 ŏ
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   QQ
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                οy
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       ó
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation-the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial
456 IIPQKDSSEETFTELNNESEEQRNFKDGEEYRIVQHEESLYGQRTKSPEENIINGSEIGV 515
                                                                                                                                                                                                                 416
                                                                   350 EGCSNEPVTFLGENDTRSIVYSNKGTNANVQEFSQEDSLAHSEPKFKDLNATSDDVWNE- 408
                                                                                                                          409 DKETDANISTSTKSEESYIAD-----YKVTRQEDWDTKK----LHQESEHANEQ--PA 455
                                                                                                                                                      332 IINKFCSAKDSEHNEKYEDTFLESEEIGTKVEVVERKEHLHTDILK-----RGSEM-- 382
                                                                                                                                                                                                                                                                                                                                                         526 SDYISIWHIQEEEIKSNSPESIANSQFSQQSSITTASTVDSKKDNGSTSFKFKPRIVSRS 685
                                                                                               DAESEPVAQKSQLDRTVCISDAGACGETLSVTSEENSLVKKKERSLSSGSNFCSEQKTSG 331
                                                                                                                                                                                                                                                                                                                            473 MAIPVLWKFLE---KYPSAEVARTADWRDVSELLKPLGLYDLR--AKTIVKFSDEYLTKQ 527
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                "Role for yeast inhibitor of apoptosis (IAP)-like proteins in cell
                                                                                                                                                                                                               ---DNNCSPTRKDFTGEKIFQEDTIPRTQIERRKTSL------
                                                                                                                                                                                                                              516 DHGEAAEVNEPLAKTSAEEHDLSSSCEDQSVSEARNKDRIEEKEVETKDENIETEKDESE
                                                                                                                                                                                                                                                                     417 YFSSKYNKEALSPPRRKAFKKWTPPRSPFNLVQETLFHDPW----KLLIATIFLNRTSGK
                                                                                                                                                                                                                                                                                                576 YHKVEENEEPEHVPLLPPLPRW------EEIQFNEPFIDENDTSNDSIDLTRSMKP
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     MEDLINE-99398681; PubMed-10468581;
Uren A.G., Beilharz T., O'Connell M.J., Bugg S.J., van Driel R.,
Vaux D.L., Lithgow T.;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           Morishita J., Matsusaka T., Yanaqida M.; "Fission yeast cut17 is required for chromosome segregation."; Submitted (AUG-1999) to the EMBL/GenBank/DDBJ databases.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              Harris D., Wood V., Rajandream M.A., Barrell B.G.;
Submitted (AUG-1998) to the EMBL/GenBank/DDBJ databases
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      Schizosaccharomyces pombe (Fission yeast).
Eukaryota; Fungi; Ascomycota; Schizosaccharomycetes;
Schizosaccharomycetales; Schizosaccharomycetaceae;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                15-JUL-1998 (Rel. 36, Created)
15-JUL-1998 (Rel. 36, Last sequence update)
20-AUG-2001 (Rel. 40, Last annotation update)
BIRI PROTEIN (CHROMOSOME SEGREGATION PROTEIN CUT17).
BIRI OR CUT17 OR SPCC962.02C.
                                        232 KG----KUTIL---KGIPIKKTKKGCRKSCSGFVQSDSKRES-
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        997 AA
                                                                                                                                                                                                                                                                                                                                                                                    528 WKY--PIELHGIGKYGNDSYRIFCVNEWKQVHP 558
                                                                                                                                                                                                                                                                                                                                                                                                                 686 RIYNPKSRVSSLNYYDNEDY-ILSNSEWNALDP 717
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        PRT;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        STANDARD;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      Schizosaccharomyces.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              SEQUENCE FROM N.A.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   SEQUENCE FROM N.A.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       CHARACTERIZATION.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  NCBI_TaxID=4896;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  STRAIN=972;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      BIR1_SCHPO
014064;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              division."
                                                                                                                                                                                                                                                                                                                                                                                                                                                           RESULT 10
           297
                                                                                                                                                                                                               383
                                                                                                                                                                                                                                                                                                                                                                                                                                                                          BIR1_SCHPO
                                                                                              272
             셤
                                                                                                                       qq
                                                                                                                                                                                                                                       g
                                                                                                                                                                                                                                                                    δ
                                                                                                                                                                                                                                                                                             g
                                                                                                                                                                                                                                                                                                                           ò
                                                                                                                                                                                                                                                                                                                                                       g
                                                                                                                                                                                                                                                                                                                                                                                                             q
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        ò
                                                                 g
                                                                                              δ
                                                                                                                                                      δ
                                                                                                                                                                                 С
                                                                                                                                                                                                            ŏ
                                                                                                                                                                                                                                                                                                                                                                                     ò
```

```
entities requires a license agreement (See http://www.isb-sib.ch/announce/
or send an email to license@isb-sib.ch).
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              33;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               73 QFGATAGTECRKSVPCGWERVVKQRLFGKTAGRFDVYFISPQGLKFRSKSSLANYLHKNG 132
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   133 ETSLK---PE---DFDFTVLSKRGIKSRYKDC-----SMAAL----TSHLONOSNN 173
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           527 ETNLKIPIPEKKRKFQEVLQSKNILVSSTEDSHEPVKVTEDSQTAIHVSKFEDLENKSME 586
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     587 SEQSLQLLSESENDDKPLIDL.IPI.LAIKRKDNLVSGVLEKGKSTSTSK---TKFDTSIVD 643
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    373 TDILKRGSEMDNNCSPTRKDFTGEKIFQEDTIPRTQIERR----KTSLYFSSKY-NKEAL 427
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    428 S---PPRRKAFKK---WTPPRSPFNLVQETLFHDPWKLLIATIFLNRTSGKMAIPVLWKFL 482
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  483 EKYPSAEVARTADWR--DVSELLK-----PLGLYDLRAKTIVKFSDEYLTKQW-KYPIEL 534
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              Gaps
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           20 VTSSERLVPDPPNDLRKEDVAMEL---ERVGEDEEQMMIKRS----SECNPLLQEPIASA 72
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                698 -NHEEANSGHTVMNVHSSLDPQPIVQPNELESGSYLKDLPDRNVGNSEKVTFQEDDINSP
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          ::|| |: :| |: :| |: || |: || |: || |: || |: || |: || |: || |: || |: || |: || |: || |: || |: || |: || |: || |: || |: || |: || |: || |: || |: || |: || |: || |: || |: || |: || |: || |: || |: || |: || |: || |: || |: || |: || |: || |: || |: || |: || |: || |: || |: || |: || |: || |: || |: || |: || |: || |: || |: || |: || |: || |: || |: || |: || |: || |: || |: || |: || |: || |: || |: || |: || |: || |: || |: || |: || |: || |: || |: || |: || |: || |: || |: || |: || |: || |: || |: || |: || |: || |: || |: || |: || |: || |: || |: || |: || |: || |: || |: || |: || |: || |: || |: || |: || |: || |: || |: || |: || |: || |: || |: || |: || |: || |: || |: || |: || |: || |: || |: || |: || |: || |: || |: || |: || |: || |: || |: || |: || |: || |: || |: || |: || |: || |: || |: || |: || |: || |: || |: || |: || |: || |: || |: || |: || |: || |: || |: || |: || |: || |: || |: || |: || |: || |: || |: || |: || |: || |: || |: || |: || |: || |: || |: || |: || |: || |: || |: || |: || |: || |: || |: || |: || |: || |: || |: || |: || |: || |: || |: || |: || |: || |: || |: || |: || |: || |: || |: || |: || |: || |: || |: || |: || |: || |: || |: || |: || |: || |: || |: || |: || |: || |: || |: || |: || |: || |: || |: || |: || |: || |: || |: || |: || |: || |: || |: || |: || |: || |: || |: || |: || |: || |: || |: || |: || |: || |: || |: || |: || |: || |: || |: || |: || |: || |: || |: || |: || |: || |: || |: || |: || |: || |: || |: || |: || |: || |: || |: || |: || |: || |: || |: || |: || |: || |: || |: || |: || |: || |: || |: || |: || |: || |: || |: || |: || |: || |: || |: || |: || |: || |: || |: || |: || |: || |: || |: || |: || |: || |: || |: || |: || |: || |: || |: || |: || |: || |: || |: || |: || |: || |: || |: || |: || |: || |: || |: || |: || |: || |: || |: || |: || |: || |: || |: || |: || |: || |: || |: || |: || |: || |: || |: || |: || |: || |: || |: || |: || |: || |: || |: || |: || |: || |: || |: || |: || |: || |: || |: || |: || |: || |: || |: || |: || |: || |: || |: || |: || |: || |: || |: || |: || |: || |: || |: || |: || |: || |: || |: || |: || |
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 475 PDMTDAATGGRVSSSSFRDKILQTNFSPRST-----IDSFSNISKKRNSEEAN--DEND
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 : | | | :: : : | | | :| 644 F----IEKPKTEIS--EVLPEEKRKAICDESQIVRVSIDRGVIKTRDVSSPVSDEKSENV
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            268 CNKADAES------EPVAQKSQLDRTVCISDA--GACGETLSVTSEENSLVKK
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 817 RTSVQNGTRSVSKNTPEKE----TKVDKIDNVSKKDVETSPGSCETSSAFAKTYAEKEVT
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   920 -----EANAPPWEPIDFSSLLESPVPNPVEPNKLSEKEL-----DMTVEQWIKF----
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             174 SNWNLRTRSKCKKD----VFMPPSSSSELQES-----RGLSNFTSTHLLLKEDEGVDD
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  ---SDSKRESV
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                313 KERSLSSGSNFCSEQKTSGIINKFCSAKDSEHNEKYEDTFLESEEIGTKVEVVERKEHLH
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              Indels 153;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 Length 997;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      952A6BAFA5C489F4 CRC64;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            535 HGIGKYGNDSYRIFCVNEWKQVHPEDHKLNKYHDWLWENHEK 576
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   964 ----MYAK-----CAKEFEEACEE--KI----EWLLEEGKR 989
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              241;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 4.3%; Score 131.5; DB 1;
20.1%; Pred. No. 0.56;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  223 VNFRKVRKPKGKVTILKGIPIKKTKKGCRKSCSGFVQ----
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            3924 AA
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              Matches 129; Conservative 119; Mismatches
                                                                                                                                                                                                                                                                                                                                                                                                                       PROSITE; PS01282; BIR_REPEAT_1; FALSE_NEG PROSITE; PS50143; BIR_REPEAT_2; 2.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   POLY-ASP.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      POLY-ASP
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   POLY-SER
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   BIR 1.
BIR 2.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            PRT;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      112579 MW;
                                                                                                                                               EMBL; AB031034; BAA83415.1; -. EMBL; AL031323; CAA20434.1; -.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     20.1%;
                                                                                                                                                                                                                                                      InterPro; IPR001370; BIR.
Pfam; PF00653; BIR; 2.
SMART; SM00238; BIR; 2.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      194
83
319
490
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         Cell division; Repeat
REPEAT 25 99
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 120
80
312
487
997 AA;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          Local Similarity
```

```
This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Blainformatics and the EMBL outstation the European Bloinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license gareement (see http://www.isb-sib.ch/announce/or send an email to license@labs.blb.ch).
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           ankyrin gene.;
Genomics 10:858-866(1991).
-- FUNCTION: ANKYRINS ATTACH INTEGRAL MEMBRANE PROTEINS TO CYTO-
-- FUNCTION: BUNDTO THE ERYTHROCYTE MEMBRANE PROTEIN
BAND 4.2, TO NA-K ATPASE, TO THE LYMPHOCYTE MEMBRANE PROTEIN
BAND TO THE CYTOSKELETAL PROTEINS FODRIN, TUBDLIN, VIMENTIN AND
DESMIN. ERYTHROCYTE ANKYRINS ALSO LINK SPECTRIN (BETA CHAIN) TO
THE CYTOPLASMIC DOMAIN OF THE ERYTHROCYTES ANION EXCHANGE PROTEIN;
THEY RETAIN MOST OR ALL OF THESE BINDING FUNCTIONS.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              TISSUE SPECIFICITY: PLASMA MEMBRANE OF NEURONS AS WELL AS GLIAL ELLS THROUGHOUT THE BRAIN.

PTM: PHOSPHORYLATED AT MULTIPLE SITES BY DIFFERENT PROFEIN KINASES AND EACH PHOSPHORYLATION EVENT REGULATES THE PROTEIN'S STRUCTURE
                                                                                                                                                                                                            MEDIINE-94075409; PubMed-8253844;
Chan W., Kordeli E., Bennett V.;
"440-kD ankyrinB: structure of the major developmentally regulated
                                                                                               Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                SEQUENCE OF 463-495 FROM N.A. MEDLINE-92009921; PubMed-183308; TSE W.T., Menninger J.C., Yang-Feng T.L., Francke U., Sahr K.E., Lux S.E., Ward D.C., Forget B.G.; "Isolation and chromosomal localization of a novel nonerythroid
                                                                                                                                                                                                                                                                                                                                                                                        Otto E., Kunimoto M., McLaughlin T., Bennett V.; "Isolation and characterization of cDNAs encoding human brain ankyrins reveal family of alternatively spliced genes."; J. Cell Biol. 114:241-253(1991).
01-OCT-1996 (Rel. 34, Last sequence update)
20-AUG-2001 (Rel. 40, Last annotation update)
ANKYRIN 2 (BRAIN ANKYRIN) (ANKYRIN B) (ANKYRIN, NONERYTHROID).
                                                                                                                                                                                                                                                                      domain and selective localization in unmyelinated axons."; J. Cell Biol. 123:1463-1473(1993).
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              Carpenter S.;
Submitted (MAY-1999) to the EMBL/GenBank/DDBJ databases
                                                                                                                                                                                                                                                                                                                                     SEQUENCE FROM N.A. (ISOFORMS 1; 2 AND 3).
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       SIMILARITY: CONTAINS 23 ANK REPEATS. SIMILARITY: CONTAINS 1 DEATH DOMAIN.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            PRODUCED BY ALTERNATIVE SPLICING.
                                                                                                                                                                                                                                                                                                                                                       TISSUE=Brain stem;
MEDLINE=91302466; PubMed=1830053;
                                                                                                                                                                             SEQUENCE FROM N.A. (ISOFORM 1).
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 EMBL; Z26634; CAB42644.1; -.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 EMBL; X56957; CAA40278.1; -. EMBL; X56958; CAA40279.2; -. EMBL; M37123; AAA62828.1; -.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           InterPro; IPR002110; ANK.
InterPro; IPR000488; Death.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         PIR; S14533; S14533.
PIR; A39643; A39643.
PIR; B39643; B39643.
PIR; S14569; S14569.
HSSP: Q00420; IAWC.
MIM; 106410;
                                                                                Homo sapiens (Human).
                                                                                                                                                                                                TISSUE=Brain stem;
                                                                                                                                    NCBI_TaxID=9606;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          REVISIONS.
```

```
31;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    3168 -----NSQMGISASTETPTKEAVSVGTKDL------PTVQTCDIPPLS--GVKQI 3209
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      3117 GATGADPUPLETSAESLALSESKETVDDEA-DLIPDSVSEEVEELPASDAQL----- 3167
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            95; Mismatches 207; Indels 206; Gaps
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           62 NPLLLQEPTASAQFGATAGTE--CRKSVPCGWERVVKQRLFGKTAGRFDVYFISPQGLKFR 119
                                                                                                                                                                                                                                                                                                                                                                                                              Q -> QFLGKLHLPTAPPPLNEGESLVSRILQLGPPGTK (IN ISOFORM 2).
MISSING (IN 1SOFORM 2 AND ISOFORM 3).
GQ -> PE (IN REF. 4).
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            2 GTTGLESLSLGDRGAAPTVTSSERLVPDPPNDLRKEDVAMELERVGEDEEQMMIKRSSEC 61
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            4.0%; Score 122; DB 1; Length 3924; 19.4%; Pred. No. 14;
                                                                                                                                                                                                                                                                                                                                                                                                                                                              > Y (IN REF. 2).
52AC496C428E29D2 CRC64;
                                                            PROSITE; PS50297; ANK_REP_REGION; 1.
PROSITE; PS50017; DEATH_DOMAIN; 1.
Cytoskeleton; Alternative splicing; Repeat; ANK repeat;
                                                                                                                                                                                                                                                                                                                                          (APPROXIMATE).
                                                                                                                                                                                                                                                                                                                                                                 (APPROXIMATE)
                                                                                                                                                                                                                                                                              REPEAT-RICH REGION.
                                                                                                                                                                                                                                                                                                                                                                                                                                              2).
                                                                                                                                                                                                                                                                                                                                                                                                                                                      QY -> HA (IN REF.
                                                                                                                                                                                                                                                                                                                                                                                                                                              I -> S (IN REF.
                                                                                                                                                                                                                                                                                                                                  REPEAT A.
                                                                                                                                                                                                                                                                                                                                                         REPEAT A.
REPEAT A (
                                                                                                                                                                                                                                                                                                           REPEAT A. REPEAT A.
                                                                                                                                                                                                                                                                                                                                                                         REPEAT A.
REPEAT A.
                                                                                                                                                                                                                                                                                                     REPEAT A.
                                                                                                                                                                                                                                                                                                                                                                                                 REPEAT A.
                                                                                                                                                                                                                                                                                                                            REPEAT A.
                                                                                                                                                                                                                                                                                                                                                   REPEAT A.
                                                                                                                                                                                                                                                                                                                                                                                         REPEAT A
                                                                                                                                                                                                                                                                                     REPEAT A
                                                                                                                                                                                                                                                                                                                                                                                                                                                               -> Y
                                                                                                                                                                                                                                                                                                                                                                                                         DEATH
                                                                                      Phosphorylation; Multigene family.
                                                                                                                    PROSITE; PS50088; ANK_REPEAT; 20
                                                                                                                                                                                                                                                                                                                                                                                                                                                                     3924 AA; 430337 MW;
      Pfam; PF00023; ank; 24.
Pfam; PF00531; death; 1.
Ffam; PF00791; 205; 1.
SMART; SM00248; ANK; 21.
SMART; SM00005; DEATH: 1.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     Best Local Similarity 19.4%
Matches 122; Conservative
InterPro; IPR000906; ZU5.
                                                                                                                                                                 393
426
426
459
525
525
591
624
657
                                                                                                                                                                                                                                                                                    1784
1796
                                                                                                                                                                                                                                                                                                                                                                                                                                              971
3582
3586
                                                                                                                                                                                                                                                                                                    808
                                                                                                                                                                                                                                                                                                                   832
                                                                                                                                                                                                                                                                                                                                                   879
891
902
                                                                                                                                                                                                                                                                                                                           844
                                                                                                                                                                                                                                                                                                                                          867
                                                                                                                                                                                                                                                                                                            1809
                                                                                                                                                                                                                                                                                                                                  1845
                                                                                                                                                                                                                                                                                                                                                   1868
                                                                                                                                                                                                                                                                                                                           1833
                                                                                                                                                                                                                                                                                                                                                         1880
                                                                                                                                                                                                                                                                                                                                                                   1892
                                                                                                                                                                                                                                                                                                                   1821
                                                                                                                                                                                                                                                                                                                                                                                                                                              CONFLICT
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             Query Match
                                                                                                                                                                                                                                                                                                                                                                                                                                                                     SEQUENCE
                                                                                                                                                                                                                                                                                                                                                                                                                 /ARSPLIC
                                                                                                                                                                                                                                                                                                                                                                                                                               VARSPLIC
                                                                                                                                                                                                                                                                                                                                                                                                                                                               CONFLICT
                                                                                                                                                                                         REPEAT
REPEAT
                                                                                                             REPEAT
                                                                                                                                                                                                                                                                              DOMAIN
                                                                                                                                                                                                                                                                                                                                                                                         REPEAT
                                                                                                                                                                                                                                                                                                                                                                                                        DOMAIN
                                                                                                     REPEAT
                                                                                                                    REPEAT
                                                                                                                           REPEAT
                                                                                                                                    REPEAT
                                                                                                                                            REPEAT
                                                                                                                                                  REPEAT
                                                                                                                                                            REPEAT
                                                                                                                                                                                                         REPEAT
                                                                                                                                                                                                                        REPEAT
                                                                                                                                                                                                                                REPEAT
                                                                                                                                                                                                                                                REPEAT
                                                                                                                                                                                                                                                       REPEAT
                                                                                                                                                                                                                                                                                                     REPEAT
                                                                                                                                                                                                                                                                                                            REPEAT
                                                                                                                                                                                                                                                                                                                           REPEAT
                                                                                                                                                                                                                                                                                                                                   REPEAT
                                                                                                                                                                                                                                                                                                                                                   REPEAT
                                                                                                                                                                                                                                                                                                                                                                   REPEAT
                                                                                                                                                                                                                                                                                                                                                                         REPEAT
                                                                                                                                                                                                                                                                                                                                                                                 REPEAT
                                                                                                                                                                                                                                                                                                                                                                                                 REPEAT
                                                                                                                                                                  REPEAT
                                                                                                                                                                           REPEAT
                                                                                                                                                                                  REPEAT
                                                                                                                                                                                                                  REPEAT
                                                                                                                                                                                                                                       REPEAT
                                                                                                                                                                                                                                                              REPEAT
                                                                                                                                                                                                                                                                      REPEAT
                                                                                                                                                                                                                                                                                     REPEAT
                                                                                                                                                                                                                                                                                             REPEAT
                                                                                                                                                                                                                                                                                                                   REPEAT
                                                                                                                                                                                                                                                                                                                                         REPEAT
                                                                                                                                                                                                                                                                                                                                                          REPEAT
 q
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           δ
```

```
3210 S-----CPDSSEPAVQVQ-LDFSTLTRSVYSDRGDDSPDSSPEEQKSVIEIPTAPMEN 3261
                                                                                                                  3262 VP-----FTESKSKIPVRTMPTSTPAPPSAEYESS------VSEDFLSSVDE 3302
                                                                                                                                                                                                       -----QLS 3329
                                                                                                                                                                                                                                                                                  3330 DLDTSV-----QKTVAPQGQDMASIAPDNRSKSESDASSLDSKTK-----CPVKTR 3375
                                                                                                                                                                                                                                                                                                                                                              SYTETETESRERAEELELESEEGATRPKILTSR--LPVKSRSTTSSCRGGTSPTKESKEH 3433
                                                                                                                                                                                                                                                                                                                                                                                                                                               3434 FFDLYRNSIEFFEEISDEASKLVDRLT----QSEREQEIVSDDESSSALEVSVIENLPPV 3489
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          3533 QERIEERLAYIADHLGFSW---TELARELDFTEEQIHQIRIENPNSLQDQSQYLLKIW-- 3587
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        "Cloning and expression of a novel human DNA binding protein, PO-GA."; Blochem. Blophys. Res. Commun. 193:779-786(1993).
120 SKSSLANYLHKNGETSLKPEDFDFTVLSKRGIKSRYKDCS-------MAALTSHLQN 169
                                                                            170 QSNNSNWNLRTRSKCKKDVFMPPSS-----SSELQESRGLSNFTSTHLLLKED--EGVDD 222
                                                                                                                                                             223 VNFRKVRKPKGKVTILKGIPIKKTKKGCRKSCSGFVQSDSKRESVCNKADAESEPVAQKS 282
                                                                                                                                                                                                                                            283 QLDRTVCISDAGACGETLSVTSEENSLVKKKERSLSSGSNFCSEQKTSGIINKFCSAK-- 340
                                                                                                                                                                                                                                                                                                                                                                                                         392 -- DFTGEKI--FQEDTIPRTQIERRKTSLYFSSKYNKEALSPPRRKAFKKWT-----PPR 442
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        443 SPFNLVQETLF--HDPWKLLIATIFLNRTSGKMAIPVLWKFLEKYP------- 486
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          ----SAEVARTAD-----WRDVSELLKPLG-----LYDLRAKTIVKFSD--EYLTKQWKY 530
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     01-0CT-1993 (Rel. 27, Created)
15-UUL-1998 (Rel. 36, Last sequence update)
20-AUG-2001 (Rel. 40, Last annotation update)
ACTIVATOR 1 140 KDA SUBUNIT (REPLICATION FACTOR C LARGE SUBUNIT) (A)
140 KDA SUBUNIT) (RF-C 140 KDA SUBUNIT)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 SEQUENCE FROM N.A., AND SEQUENCE OF 469-480; 571-580 AND 677-699. MEDLINE-94068535; PubMed-8248204; Bunz F., Kobayashi R., Stillman B., "Chbayas encoding the large subunit of human replication factor C."; Proc. Natl. Acad. Sci. U.S.A. 90:11014-11018(1993).
                                                                                                                                                                                                                                                                                                                        341 ----DSEHNEKYEDTFLESEEIGTKVEVVERKEHLHTDILKRGSEMDNNCSPTRK----
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
                                                                                                                                                                                                 3303 ENKADEAKPKSKLPV--KVPLQRVEQ-----
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            Rajavashisth T.B., Tripathi S.;
Submitted (FEB-1998) to the EMBL/GenBank/DDBJ databases.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          PRT; 1147 AA.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             3588 --- LERDGKHATDTNLVECLTKINRMDIVH 3614
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         531 PIELHGIGKYGNDSYRIFC---VNEWKQVH 557
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      FUNCTION, AND INTERACTION WITH PCNA.
MEDLINE=97153138; PubMed=8999859;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  MEDLINE=93290676; Pubmed=8512577;
Lu Y., Zeft A.S., Riegel A.T.;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           (DNA-BINDING PROTEIN PO-GA).
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     STANDARD;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      Homo sapiens (Human)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    SEQUENCE FROM N.A.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               SEQUENCE FROM N.A.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             TISSUE=Hepatoma;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             NCBI_TaxID=9606;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 OR RFC140
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   AC15_HUMAN
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        487
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               RFC1
ò
                                    g
                                                                                                                                                                                                                                                                                q
                                                                                                                                                                                                                                                                                                                                                                                                                                             q
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 g
                                                                            ò
                                                                                                                q
                                                                                                                                                             ò
                                                                                                                                                                                                 g
                                                                                                                                                                                                                                         ò
                                                                                                                                                                                                                                                                                                                        ò
                                                                                                                                                                                                                                                                                                                                                              q
                                                                                                                                                                                                                                                                                                                                                                                                     δ
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        ŏ
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              q
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         ô
```

```
This SWISS-PROT entry is copyright. It is produced through a collaboration
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               between the Swiss Institute of Bioinformatics and the EMBL outstation. The European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See http://www.isb-sib.ch/announce/or send an email to license@isb-sib.ch).
                                                                                                                                           PCNA AND ACTIVATOR 1. THE 140 SUBUNIT BINDS TO THE PRIMER-TEMPLATE JUNCTION. BINDS THE PO-B TRANSCRIPTION ELEMENT AS WELL AS OTHER GA RICH DAM SEQUENCES. COULD PLAY A ROLE IN DMA TRANSCRIPTION REGULATION AS WELL AS DNA REPLICATION AND/OR REPAIR. CAN BIND
                                                                                                                                                                                                                                                                                            SINGLE OR DOUBLE-STRANDED DNA.
FUNCTION: INTERACTS WITH C-TERMINUS OF PCNA. 5' PHOSPHATE RESIDUE
IS REQUIRED FOR BINDING OF THE N-TERMINAL DNA-BINDING DOMAIN TO
DUPLEX DNA, SUGGESTING A ROLE IN RECOGNITION OF NON-PRIMER
TEMPLATE DNA STRUCTURES DURING REPLICATION AND/OR REPAIR.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 ATP (POTENTIAL).
C2HC-TYPE (POTENTIAL).
NUCLEAR LOCALIZATION SIGNAL (POTENTIAL).
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              PROSITE; PS50172; BRCT; 1.
DNA replication; ATP-binding; Transcription regulation; DNA-binding;
                                                                                                                                                                                                                                                                                                                                                                              SUBUNIT: HETEROPENTAMER OF SUBUNITS OF 140/145, 40, 38, 37, AND 36.5 KDA THAT FORMS A COMPLEX WITH PCNA IN THE PRESENCE OF ATP. SUBCELLULAR LOCATION: NUCLEAR.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                SIMILARITY: BELONGS TO THE ACTIVATOR 1 140 KDA SUBUNIT FAMILY. SIMILARITY: CONTAINS 1 BRCT DOMAIN.
                                                                                                                                                                                                                                                                                                                                                                                                                                TISSUE SPECIFICITY: WIDE TISSUE DISTRIBUTION. UNDETECTABLE IN
Mossi R., Jonsson Z.O., Allen B.L., Hardin S.H., Huebscher U
"Replication factor C interacts with the C-terminal side of
proliferating cell nuclear antigen.";
J. Biol. Chem. 272:1769-1776(1997).
                                                                                              MEDLINE-98371221; ubbMed-9705493;
Allen B.L., Uhlmann F., Gaur L.K., Mulder B.A., Posey K.L.,
Jones L.B., Hardin S.H.;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    5BC2878FDD2496D9 CRC64;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  E -> K (IN REF. 1).
5 -> R (IN REF. 1).
5 -> AA (IN REF. 1).
N -> G (IN REF. 1).
N -> G (IN REF. 1).
A -> S (IN REF. 1).
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   InterPro; IPR001939; AAA_subfam.
InterPro; IPR001357; BRCT.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    1147 AA; 128282 MW;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 EMBL; L14922; AAA86853.1; -. EMBL; L23320; AAA16121.1; -. EMBL; Z22642; CAA80355.1; -.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             EMBL; AF040250; AAB99788.1;
PIR; JN0599; JN0599.
PIR; S32762; S32762.
MIM; 102579; -.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  Activator; Nuclear protein.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    SMART; SM00292; BRCT; 1
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 657
766
1124
326
613
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 Pfam; PF00004; AAA; 1.
Pfam; PF00533; BRCT; 1.
                                                                                                                                                                                                                                                                                                                                                                                                                                                 PLACENTAL TISSUE.
                                                                               DNA-BINDING ACTIVITY
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     IPR000862;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   650
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   749
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     402
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  CONFLICT
CONFLICT
SEQUENCE
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     InterPro;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  CONFLICT
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 NP_BIND
ZN_FING
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   CONFLICT
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      CONFLICT
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    DOMAIN
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    DOMAIN
```

Indels 259; Gaps

Length 1147;

111 ISPQGLKFRSKSSLANYLHKNGETSLKPEDFDFTVLSKRGIK-----SRYKDCSMAALT 164

3.9%; Score 120.5; DB 1; 9.4%; Pred. No. 3.4; ve 86; Mismatches 194;

Local Similarity 19.4%;

Conservative

Matches 130;

ô

Query Match

qq	7	:
Qy		- 20
2 2	201	SKRKKIIIUSDSESEEFILOVKNARKFPEKLEVSSRFGKISKUDVIIISEIDEEDDFMCK 119 SRGISNFGTHIILKEDEGUDDVNFRKVRKPKRKVTIIKGIPIKKTKKGR 251
G 20		ASKSKENGRSTNSHLGTSNMKKNEENTKTKNKPLSPIKLTPTSVL 1
QY	252 F	KSC-SGFVQSDSKRESVCNKADAESEPVAQKSQLDRTVCISDAGACG 297
qq	167	DYFGTGSVORSNKKMVASKRKELSONTDESGLNDEATAKOLOLDEDAELERQLHEDEEFA 226
oy de	298 E	ETLSVISEENSLVKKKERSLSSGSNFCSEQ
3 8		-BUNDEV FORDE TEGEDICAN WINDERFULL HARMEN FORDER 38
2 A	00	ANY DESKRESQQHSKSSADKIGEVSSPKASSKLAIMKRKEESSYKEIEPVASKRENAI 34 SKYESSKESQQHSKSSADKIGEVSSPKASSKLAIMKRKEESSYKEIEPVASKRKENAI 34
٥y	387	ALSPPR
Dp	346	KLKGETKTPKKTKSSPAKKESVSPEDSEKKRTNYQAYRSYLNREG 390
ογ P	432 F	RKAFKKWTPPRSPFNLVQETLFHDPWKLLIATIFL 466
3	1	
Qy Db	451 (NRTSGKMAIPVLWKFLE
Qy	503	
Db	511	: : : VQCKRKISPSKKESESKKSRPTSKRDSLAKTIKKETDVFWKSLDFKEQVAEETSGD 566
δy	530	
Ωp	267	SKARNLADDSSENKVENLLMVDKYKPTSLKTITGQQGDQSCANKLLLMLRNWQKSS 622
οy	558	PEDHKLNKY 566
qq	623	: : SEDKKHSKF 631
RESULT YFI6_YI	LT 13 _YEAST	
a z	YFI6_YEAST	YEAST STANDARD; PRT; 1233 AA.
점	01-NO	Rel.
DT DE	01 - NO 01 - NO HYPOTI	01-NOV-1995 (Rel. 32, Last sequence update) 01-NOV-1995 (Rel. 32, Last annotation update) HYPOTHETICAL 137.7 KDA PROTEIN IN UGS1-FAB1 INTERGENIC REGION.
	YFR016C Sacchar	isiae (Baker's yeast).
	Eukar	<pre>ji; Ascomycota; les; Saccharomyc</pre>
	NCBI_	TaxID=4932;
RC S	SEQUE	SEQUENCE FROM N.A. STRAIN-S288C / AB972;
8 8 8	Murak	mbDuine_9340023; Fubmec-101403; Murakami Y., Naitou M., Hagiwara H., Shibata T., Ozawa M., Sasanuma S.I., Sasanuma M., Tsuchiva Y., Soeda E., Yokovama K.,
RA	Yamaz	Eki T.;
RT	Sacch	(566)
888	S - i -	Η,

```
This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the BMBL outstation the European Bioinformatics institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See http://www.isb-sib.ch/announce/or send an email to license@lsb-sib.ch).
                                                                                                                                                                                   972 INISDEFÇRTVELPELEKQDIKDNKGEDKELEVEETEKETSLPDLVVEENITEEKNEIKQ 1031
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              1032 EEEEVSQLDFNETESISKEAPNNDENGFEDQSTRENPKKASADDIFKDILDETN----- 1085
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  | : | | : : | | : : | | 853 ENEDVEVDTEEDAEVENSERTEFIKVKAELGNLDAPKEAEVTAELNKENEDVEVAATSKE 912
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        373 ---TDILKRGSEMDNNCSPTRKDFTG-----EKIFQEDTIPRTQIERRKTSLYFSSKY 422
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           423 NKEALSP---PRRKAFKKWTPPRSPFNLVQETLFHDPWKLLIATIF---LNRTSGKMAIP 476
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               79 ----GTECRKSVPCGWERVVKQRLFGKTAGRFDVYFISPQGLKFRS--------- 120
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 658 ESAEGVEVEKSKTPESPKVVKR----CTSGR-----PEDLQINERDPEVLKEDVRVPD 706
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               121 ---KSSLANYLHKNGE-----KPEDFDFTV 145
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     265 ESVCNKADAESEPVAQKSQLDRTVCIS-----DAGACGETLSVTSEENSLVKKKERSLS 318
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     319 SGSNFCSEQKTSGIINKFC----SAKDSEHNEKYEDTFLESEEIGTKVEVVERKEHLH- 372
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        913 DIETKCSEPAETPIEDGTCTEAEVSKKDAEAVTK-EDENMENSKIAEALKDVTGDQEIDD 971
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         546 DNLVLEDEAEAPTQENKPTEVVGEIDIPDAPRDDVEIVEAVEKNIIPEDLEVAKEDQEGE 605
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            606 QVK------LDEPVKAMKDDKIAMRGAESISEDMKKKQEGTAELSNEKAKKEVDETAR 657
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  146 LSKRGIKSRYKDCSMAALTSHLQNQSNN-SNWNLRTRSKCKKDVFMPPSSSSELQESRGL 204
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             766 -- OEGDKITGKD----TNHEHGEATEAASENSKASDVGTAEKYIEPSSESVKKDT--- 814
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  205 SNFTSTHLLLKEDEGVDDVNFRKVRKPKGKVTILKGIPIKKTKKGCRKSCSGFVQSDSKR 264
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  815 -------VTAELNK 852
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  01-OCT-1996 (Rel. 34, Last sequence update)
20-AdG-2001 (Rel. 40, Last annotation update)
ACTIVATOR 1 140 KDA SUBUNIT (REPLICATION FACTOR C LARGE SUBUNIT) (A1 140 KDA SUBUNIT) (RF-C 144 KDA SUBUNIT) (ACTIVATOR 1 LARGE SUBUNIT) (A1-P145) (DIFFERENTIATION SPECIFIC ELEMENT BINDING PROTEIN)
                                                                                                                                                                                                                                                                                                                                                                                                           Gaps
                                                                                                                                                                                                                                                                                                                                                                                                                                                           7 ESLSLGDRGAAPTVTS-----SERLVPDPPND-----LRKEDVAMELERVGEDEEQM 53
                                                                                                                                                                                                                                                                                                                                                        3.9%; Score 120.5; DB 1; Length 1233; 18.9%; Pred. No. 3.7; tive 91; Mismatches 228; Indels 199;
                                                                                                                                                                                                                                                           Hypothetical protein.
SEQUENCE 1233 AA; 137697 MW; C8A7CD2C6F0892F6 CRC64;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            477 VLWKFLEKYPSAEVARTADWRDVSELLKPLGLYDLRAKT 515
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            PRT; 1131 AA.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            01-JUN-1994 (Rel. 29, Created)
                                                                                                                                                                                                        EMBL; D50617; BAA09255.1; -.
                                                                                                                                                                                                                                                                                                                                                                                                           Matches 121; Conservative
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 54 MIKRSSECNPLLQEPI----
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               STANDARD;
                                                                                                                                                                                                                                     SGD; S0001912; YFR016C.
                                                                                                                                                                                                                                                                                                                                                                                      Best Local Similarity
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               AC.15_MOUSE
                                                                                                                                                                                                                                                                                   SEQUENCE
                                                                                                                                                                                                                                                                                                                                                              Query Match
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    P35601
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       AC15_MOUSE
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            ID DT DE DE DE DE DE DE
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  qq
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                QQ
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        Ωp
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           QQ
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              QY
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           q
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              δŏ
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              qq
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            QQ
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               qq
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  qq
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  δ
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     δò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     QY
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     δλ
             δŏ
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 ŏ
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    οy
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    ŏ
```

```
This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the BWBL outstation the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See http://www.isb-sib.ch/announce/or send an email to license@isb-sib.ch).
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  MEDINE-90229765; Pubmed-1691767;
MEDINE-90229765; Pubmed-1691767;
MEDINE-90229765; Pubmed-1691767;
MEDINE-90229765; Pubmed-1691767;
MEVALUATION of Inter- and intramolecular primary structure homologies of interferons by a Monte Carlo method.";
J. Interferon Res. 10:31-31(1990).
--- PUNCTION: THE ELONGATION OF PRIMED DNA TEMPLATES BY DNA POLYMERASE PEURA AND MEDICAN REQUIRES THE ACTION OF THE ACCESSORY PROTEINS
PRIME AND MEDICAN REQUIRES THE ACTION OF THE PRIMER-
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     "Differentiation-specific element binding protein (DSEB) binds to a defined element in the promoter of the angiotensinggen gene required for the irreversible induction of gene expression during differentiation of 373-L1 adipoblasts to adipocytes."; Mol. Endocrinol. 9:487-501(1995).
                                                                                                                                                                                                                                                                                                                                                                                             Burbelo P.D., Utani A., Pan Z., Yamada Y.; "Cloning of the large subunit of activator 1 (replication factor C) reveals homology with bacterial DNA ligases."; Proc. Natl. Acad. Sci. U.S.A. 90:11543*11547(1993).
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            -!- SUBUNIT: HETEROPENTAMER OF SUBUNITS OF 140/145, 40, 38, 37, AND 36-5 KDA THAT FORMS A COMPLEX WITH PCNA IN THE PRESENCE OF ATP.-!- SUBCELLULAR LOCATION: NUCLEAR.-!- SUBLILARITY: BELONGS TO THE ACTIVATOR 1 140 KDA SUBUNIT FAMILY.-!- SIMILARITY: CONTAINS 1 BRCT DOMAIN.
                                                                                                         Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Musinae; Mus
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        Gordon D.F.;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       Luckow B., Bunz F., Stillman B., Lichter P., Schuetz G.; "Cloning, expression, and chromosomal localization of the 140-kilodalton subunit of replication factor C from mice
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              STRAIN=LAF1;
Lossie A.C., Haugen B.H., Wood W.M., Camper S.A., Gordor
Submitted (OCT-1994) to the EMBL/GenBank/DDBJ databases.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       Submitted (FEB-1994) to the EMBL/GenBank/DDBJ databases
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                Mol. Cell. Biol. 14:1626-1634(1994)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         MEDLINE=95388065; PubMed=7659092; McGehee Habener J.F.;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         MEDLINE=94158835; PubMed=8114700;
                                                                                                                                                                                                                                                                                                                     STRAIN=BALB/C;
MEDLINE=94089669; PubMed=8265586;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           EMBL; U01222; AAA21643.1; -. EMBL; X72711; CAA51260.1; -.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        EMBL; U36441; AAA79698.1; -. EMBL; U07157; AAC52140.1; -. EMBL; U15037; AAB60452.1; -.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       SEQUENCE OF 1-565 FROM N.A.
(ISRE-BINDING PROTEIN).
                                 RFC1 OR RECC1 OR IBF-1.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              TEMPLATE JUNCTION.
                                                                                musculus (Mouse).
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               [3]
SEQUENCE FROM N.A.
                                                                                                                                                                                                                                                                               SEQUENCE FROM N.A.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      SEQUENCE FROM N.A.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             SEQUENCE FROM N.A.
                                                                                                                                                                                                  NCBI_TaxID=10090;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               STRAIN=SWISS
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       TISSUE=Liver
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         Hadue S.J.;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              numans.
```

```
405
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   454 HDPWKLLIATIFLNRTSGKMAIPVLWKFLEKYPSAEVARTADWRDVSELLKPLGLYDLRA 513
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      203 GLSNFTSTHLLLKEDEGV-----DDVNFRKV-----RKPKGKVTILKGIPIKKTKKGCRK 252
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 253 SCSGF--VQSD---SKRESVCNKAD--------AESEPVAQ--KSQLDRTV 288
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             -----VEVVERKEHLHT 373
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                367 EDSEKRRTNYQAYRSYLNREGPKALGSKEIPKGAENCLEGLTFVITGVLESIERDE-AKS 425
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       406 -----RTQIERRKTSLYFSSKYN---KEALSPPRR----KAFKKWTPPRSPFNLVQETLF 453
                                                                                                                                                                                                                                                                                                                                                                                                                                                    119 RSKSSLANYLHKNGETSLK-----PEDFDFTVLSKRGIKSRYKDCSMAALTSHLQN 169
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             247 GEESFSSVQDDLSKAEKOKSPNKAELFSTARKTYSPAKHGKGRASEDAKOPCKSAHRKEA 306
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                82 KKKSEKLSLSYKPGKVSQKDPVTYVSETDEDDDF-VCKKAASKSKENGVS----TNSYLG 136
                                                                                                                                                                                                                                                                                                                                                                                                                      76; Mismatches 182; Indels 218; Gaps
                                      Pfam: PF00004; AAA: 1.
Pfam: PF0053; BRCT: 1.
SMAFT: SMAFT: 1.
SMAFT: SMO529; BRCT: 1.
DNA replication; ATP-binding; Transcription regulation; DNA-binding;
                                                                                                                                                                           NUCLEAR LOCALIZATION SIGNAL (POTENTIAL)
                                                                                                                                                                                          INTERFERON - STIMULATED - RESPONSE - ELEMENT
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          486 YEMAAEAEMKKEKSKLERTPQKNDQGKRKISPAKKESESKKCKLTLLKNSPMKAVKKEAS
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             ----SSSELQESR
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                197 LNDEAIAKOLOLDEDAELEROLHEDEEFARTLALLDEEPK------IKKARKDSEE
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            289 CISDAGACGETLSVTSEENS-----LVKKKKERSLSSGSNFCSEQKT--SGIINKFCSA
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              374 DILKRGSEMDNNCS-----TIP----
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           426 LIERYGGKVTGNVSKKTNYLVMGRDSGQSKSDKAAALGTKILDEDGLLDLIRTMPGKRSK
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               ------KPASLKNI--
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                514 KTIVKESDEYLTKQWKYPIELHGIGKYGNDSYRIFCVNE------WKQVHPEDHK 562
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                ---IGQQGDQS----CANKLLRWLRNWHKSSPEEKK 611
                                                                                                                                                                                                                     Y -> N (IN REF. 3).
E -> EPDFCLSCLIFFGIQ (IN REF. 4)
V -> A (IN REF. 5).
N -> S (IN REF. 4).
MISSING (IN REF. 3 AND 4).
S -> N (IN REF. 1).
T -> A (IN REF. 3).
K -> KQ (IN REF. 3).
                                                                                                                                                                                                                                                                                                                                                                                         DB 1; Length 1131;
                                                                                                                                                              C2HC-TYPE (POTENTIAL)
                                                                                                                                                ATP (BY SIMILARITY)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  340 KDSEH---NEKYEDTFLESE---EIGTK-----
                                                                                                                                                                                                            BINDING REGION
                                                                                                                                                                                                                                                                                                                                                                                         3.9%; Score 120;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               170 OSN--NSNWNLRTRSKCKKDVFMPPS-----
                                                                                                                                                                                                                                                                                                                                                                                                       Pred. No.
Interpro; IPR001939; AAA_subfam.
Interpro; IPR001357; BRCT.
                                                                                                                                                                                                                                                                                                                                             X
X
                                                                                                                                                                                                                                                                                                                                             125984
                                                                                                                                                                                                                                                                                                                                                                                                       20.1%;
                                                                                                                    Activator; Nuclear protein.
                                                                                                                                                                                                                                                                                                                                                                                                                        Conservative
                            IPR000862; RFC
                                                                                                                                                              751
1108
528
                                                                                                                                                                                                                        66
187
254
559
614
                                                                                                                                                642
                                                                                                                                                                                                                                                                                                                               1104
                                                                                                                                                                                                                                                                                                                                              AA;
                                                                                                                                                                                                                                                                                                                                                                                         Query Match
Best Local Similarity
                                                                                                                                                                                                                                                                    559
614
945
1071
11131
                                                                                                                                                                                                                        66
187
254
                                                                                                                                                              734
1104
354
                                                                                                                                                                                                                                                                                                                                                                                                                        120;
              InterPro;
InterPro;
                                                                                                                                                                                                                                                                                                CONFLICT
CONFLICT
CONFLICT
                                                                                                                                                                                                                                                                    CONFLICT
                                                                                                                                                                                                                                                                                                                                             SEQUENCE
                                                                                                                                               NP_BIND
ZN_FING
DOMAIN
                                                                                                                                                                                                                          CONFLICT
                                                                                                                                                                                                                                        CONFLICT
                                                                                                                                                                                                                                                       CONFLICT
                                                                                                                                 DOMAIN
                                                                                                                                                                                            DOMAIN
                                                                                                                                                                                                                                                                                                                                                                                                                         Matches
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                Dp
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                qq
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               ολ
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           qc
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     Dp
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              qq
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            ŏ
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               QΫ́
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    ŏ
```

RESULT 15 YM67_YEAST

MGD; MGI:97891; Reccl.

30;

```
between the Swiss Institute of Bioinformatics and though a contraction the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See http://www.isb-sib.ch/announce/or send an email to license@isb-sib.ch).
                                                                                                                                                                                                                                                                                                                                                                                                                                                                             This SWISS-PROT entry is copyright. It is produced through a collaboration
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        3.9%; Score 120; DB 1; Length 1658;
19.2%; Pred. No. 5.9;
Live 69; Mismatches 148; Indels 116; Gaps
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  971 IVEPENEKVNKVHEEETLFEANVSSSVNVQNKDMHTDVINQEAQANYEAGERKYYIQNTD 1030
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       Skelton J., Churcher C.M., Barrell B.G., Rajandream M.A., Walsh S.V.; Submitted (JUN-1995) to the EMBL/GenBank/DDBJ databases.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      928 VVDKL-----KGTKKDTLG 970
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  825 SAINYTNVTGDSSC--EDIIETASNVEENLRYCEKDMNEAEMSSGDECVKONDDGS---- 878
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      879 ---KTQISFSTDSPDNFQESNDNTEFSSTKYKVRNSDLEDDESLKK-----ELTKAE 927
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             302 VTSEENSLVKKK-----ERSLSSGSNFCSEQKTSGIINKFCSAKDSEHNEKYEDTFLE 354
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           355 SEE------IGTKVEVVERK--EHLHTDIL--KRGSEMDNN------ 385
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              386 ----CSPTRKDFTGEKIFQEDTIPRTQIERRKTSLYFSSKYNK-----EALSPPRR 432
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           123 SLANYLHKNGETSLKPEDFDFTVLS-KRGIKSRYKDCSMAALTSHLQNQSNNSNWNLRTR 181
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           182 SKCKKDVFMPPSSSSELQESRGLSNFTSTHLLLKEDEGVDDVNFRKVRKPKGKVTILKGI 241
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            242 PIKKTKKGCRKSCSGFVQSDSKRESVCNKADAESEPVAQKSQLDRIVCISDAGACGETLS 301
                                                                                                                                           Saccharomyces cerevisiae (Baker's yeast).
Eukaryota; Fungi; Ascomycota; Saccharomycotina; Saccharomycetes;
Saccharomycetales; Saccharomycetaceae; Saccharomyces.
          003661; Q04988;

01-NOV-1997 (Rel. 35, Created)

01-NOV-1997 (Rel. 35, Last sequence update)

01-NOV-1997 (Rel. 35, Last annotation update)

HYPOTHETICAL 187.1 KDA PROTEIN IN GUAL-ERG8 INTERGENIC REGION.

YMR219W OR YM8261.13 OR YM9959.01.
                                                                                                                                                                                                                                                                                         Dedman K., Brown D., Bowman S., Barrell B.G., Rajandream M.A., Walsh S.V.;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    ----PESD----RSNIFSSPIRVIGAVV---KGVGKV-VDVAESFVKK 1170
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            433 KAFKKWTPPRSPFNLVQETLFHDPWKLLIATIFLNRTSGKMAIPVLWKFLEK 484
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       Hypothetical protein. SEQUENCE 1658 AA; 187137 MW; 3893F968305A757D CRC64;
                                                                                                                                                                                                                                                                                                                                  Submitted (JUN-1995) to the EMBL/GenBank/DDBJ databases.
  PRT; 1658 AA.
                                                                                                                                                                                                                                                                                                                                                                       SEQUENCE OF 608-1648 FROM N.A.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             EMBL; Z49809; CAA89934.1; -.
                                                                                                                                                                                                                                               SEQUENCE OF 1-711 FROM N.A.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         Ouery Match
Best Local Similarity 19.2%
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                ; Z49939; CAA90190.1;
S0004832; YMR219W.
  STANDARD;
                                                                                                                                                                                                                                                                       STRAIN=S288C / AB972;
                                                                                                                                                                                                                                                                                                                                                                                             STRAIN=S288C / AB972;
                                                                                                                                                                                                            NCBI_TaxID=4932;
                                                                                                                                                                                                                                                                                                              Walsh
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     EMBL;
q
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              δλ
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           Ω
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           δŏ
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            δ
```

Search completed: January 15, 2002, 14:07:26

Job time: 955 sec

```
GenCore version 4.5
Copyright (c) 1993 - 2000 Compugen Ltd.
```

OM protein - protein search, using sw model

January 15, 2002, 14:06:25; Search time 134.53 Seconds (without alignments) 344.669 Million cell updates/sec Run on:

Title: Perfect score:

US-09-724-296-37 1662 1 MIFRFGFVSNAMSLWDASPA......ELSSIRGVRRIGGGALQWKS 317 Sequence:

BLOSUM62 Gapop 10.0 , Gapext 0.5 Scoring table:

473505 seqs, 146272329 residues Searched:

473505 Total number of hits satisfying chosen parameters:

Minimum DB seq length: 0 Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0% Maximum Match 100% Listing first 45 summaries

SPTREMBL_17:* Database :

sp_archea:*
sp_bacteria:*
sp_fungi:*
sp_human:*
sp_lnvertebrate:*
sp_nammal:* sp_virus:*
sp_vertebrate:*
sp_unclassified:* sp_organelle:*
sp_phage:*
sp_plant:*
sp_rodent:* sp_mhc:* 10: 111: 12: 13:

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

STIMMARTES

S		Description	0914f0 bacillus ce	0914e9 bacillus ce							09x2v9 bacillus an	09p977 candida alb	09vzs7 drosophila	O9p411 candida alb	09hft8 candida alb			_		O9r095 rattus norv	Ol4014 schizosacch
SUMMARIES		ID	Q9L4F0	Q9L4E9	Q9K9P8	010988	001408	6W0S60	Q9RTE6	Q9HGM8	Q9X2Y9	Q9P977	Q9V2S7	Q9P411	Q9HFT8	P95930	Q9KB81	045152	Q9NGC2	Q9R095	014014
		DB	7	7	7	٣	m	7	7	٣	7	٣	S	m	m	-	~	7	ß	11	3
		Match Length DB	317	325	322	599	929	305	326	695	315	1690	2272	1690	1690	283	823	331	505	1744	696
	* 2	Match	60.2	60.2	29.5	21.6	20.8	17.7	17.7	9.9	6.3	6.3	6.3	6.1	6.1	6.1	9.0	5.9	5.9	5.9	5.8
		Score	1001	1001	484.5	359.5	345	293.5	293.5	109.5	104.5	104.5	104	101.5	101.5	101	100	86	98	97.5	96.5
	Result	NO.		7	e	4	S	9	7	8	σ	10	11	12	13	14	15	16	17	18	19

Q9vlt3 drosophila O45519 caenorhabdi O19702 caenorhabdi O04146 acetabulari	V91723 coenorhabdi 017235 caenorhabdi 09tzx4 caenorhabdi 020833 caenorhabdi	075226 homo sapien 075226 homo sapien 09ubu7 homo sapien 09upy9 maize strea	Varyo Prococus Objzz7 neisseria m Objcn0 xylella fas Objr10 caenorhabdi Obj2f7 homo sapien Objzf7 homo sapien	008461 drosophila 09sthl arabidopsis 09k940 bacillus ha 09zvt9 arabidopsis 017470 caenorhabdi 018125 caenorhabdi 019881 caenorhabdi 044733 caenorhabdi
09VLT3 045519 019702 0 004146		1 075226 1 090807 12 09YPX9	09577 09577 093719 0942F7 096RA1	00957H1 0957H1 0957H9 0 0927T9 011125 019881
1760 5 775 5 553 5 1145 1	805 5 1066 5 674 5		921 23 23 24 25 25 25 25 25 25 25 25 25 25 25 25 25	1409 558 1020 1020 1020 1020 1020 1020 1020 102
55.7	~ w w w w	2000	ומממממ	សលលលលលល សល់សស្444
96.5 94.5 94.5	93.55 93.55 93.55 5	900 92	91.5 91.5 91.5 91.5	9 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0
20 21 23 23	7 5 5 5 6 5 1 6 7 8 8 8 9 7 8 9 7 8 9 7 8 9 9 8 9 9 9 9	30 30 30 30 30 30 30 30 30 30 30 30 30 3	1 6 8 8 8 8 8 8 8 8 8 8 8 8 8 8 8 8 8 8	338 4444 5444 5443 5443

ALIGNMENTS

RESULT TO THE PRESULT	SULT 1 0404F0 PRELIMINARY; PRT; 317 AA. 0914F0 OUT-2000 (TrEMBLrel. 15, Created) 01-0CT-2000 (TrEMBLrel. 15, Last sequence update) 01-0CT-2000 (TrEMBLrel. 15, Last sequence update) 01-0CT-2000 (TrEMBLrel. 16, Last annotation update) 01-0CT-2000 (TrEMBLrel. 16, Last sequence update) 02-0CT-2000 (TrEMBLrel. 16, Last sequence update) 03-0CT-2000 (TrEMBLrel. 16, Last sequence update) 04-0CT-2000 (TrEMBLRELSESTOPLATHODVAMDEVTERRERELICELXTHOLRTSFHPNOFTLFTSD 120 04-0CT-2000 (TrEMBLRELEGATIONICAL UPDATE UPDA
--	--

7

```
7
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         SEGUENCE FROM N.A.
SERAIN=ATCC 14579 TYPE STRAIN;
MEDLINE=AUG5537; PubMed=10589720;
Okstad O., Gominet M., Purnelle B., Rose M., Lereclus D., Kolsto A.B.;
"Sequence analysis of three Bacillus cereus loci under PICR-regulated genes encoding degradative enzymes and enterotoxin.";
Microbiology 145:3129-3138 [1999).
EMBL; AJ243712; CAB69814.1;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                61 HGIPLYRFSSSIVPLATHPDVMWDFVTPFQKEFREIGELVKTHQLRTSFHPNQFTLFTSP 120
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         121 KESVTKNAVTDMAYHYRMLEAMGIADRSVINIHIGGAYGNKDTATAQFHQNIKQLPQEIK 180
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         181 ERMILENDDKTYTTEETLQVCEQEDVPFVFDFHHFYANPDDHADLNVALPRMIKTWERIG 240
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 189 KQMTLENDDKTYTTAETLSICQKENIPFVFDYHHHMANLCEE-PLEELLPAIFETWSHTN 247
181 ERMTLENDDKTYTTEETLQVCEQEDVPFVFDFHHFYANPDDHADLNVALPRMIKTWERIG 240
                                         241 LQPKVHLSSPKSEQAIRSHADYVDANFL---LERFRQWGTNIDFMIEAKQKDKALLRLMD 297
                                                                                                                                                                                            240 ISPKVHISSPRSEKEFRAHAEYIDLEFIKPFLHIAKKHNHNFDIMIESKQKDLALFQLID 299
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                241 LQPKVHLSSPKSEQAIRSHADYVDANFL---LERFRQWGTNIDFMIEAKQKDKALLRLMD 297
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         248 ISPKVHISSPRSEKEFRAHAEYIDLEFIKPFLHIAKKHNHNFDIMIESKQKDLALFQLID 307
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   4; Gaps
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        1 MIFRFGFVSNAMSLWDASPAKTLTFARYSKLSKTERKEALLTVTKANLRNTMRTLHYIIG 60
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  9 MLIREGYVSHAMALWDCSPAKTMTFTSFKKLSKOEREDKLYHVIRONLEHTIRILHYNIA 68
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            60.2%; Score 1001; DB 2; Length 325; 58.2%; Pred. No. 2.8e-74; ive 55; Mismatches 74; Indels
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  325 AA; 37900 MW; 07EFBDD1FA580653 CRC64;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   01-0CT-2000 (TrEMBLrel. 15, Created)
01-0CT-2000 (TrEMBLrel. 15, Last sequence update)
01-MAR-2001 (TrEMBLrel. 16, Last annotation update)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     Bacteria; Firmicutes; Bacillus/Clostridium group;
Bacillus/Staphylococcus group; Bacillus.
NCBL_TaxID=1396;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              322 AA.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      325 AA
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           Q9K9P8;
01-OCT-2000 (TrEMBLrel. 15, Created)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          PRT;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      PRT;
                                                                                                                                                                                                                                                                                                                                                    300 ELSAIRGIKRISGAMLQW 317
                                                                                                                                                                                                                                                                                                       298 ELSSIRGVKRIGGGALQW 315
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   298 ELSSIRGVKRIGGGALQW 315
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            308 ELSAIRGIKRISGAMLQW 325
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          Matches 185; Conservative
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               PUTATIVE UV-ENDONUCLEASE.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              PRELIMINARY;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      PRELIMINARY;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     Query Match
Best Local Similarity
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             Bacillus cereus.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     Endonuclease
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      SEQUENCE
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       Q9L4E9;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          Q9K9P8
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      Q9L4E9
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                ٣
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            ~
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                RESULT
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         091469
091410
091400
091400
091400
091400
091400
091400
091400
091400
091400
091400
091400
091400
091400
091400
091400
091400
091400
091400
091400
091400
091400
091400
091400
091400
091400
091400
091400
091400
091400
091400
091400
091400
091400
091400
091400
091400
091400
091400
091400
091400
091400
091400
091400
091400
091400
091400
091400
091400
091400
091400
091400
091400
091400
091400
091400
091400
091400
091400
091400
091400
091400
091400
091400
091400
091400
091400
091400
091400
091400
091400
091400
091400
091400
091400
091400
091400
091400
091400
091400
091400
091400
091400
091400
091400
091400
091400
091400
091400
091400
091400
091400
091400
091400
091400
091400
091400
091400
091400
091400
091400
091400
091400
091400
091400
091400
091400
091400
091400
091400
091400
091400
091400
091400
091400
091400
091400
091400
091400
091400
091400
091400
091400
091400
091400
091400
091400
091400
091400
091400
091400
091400
091400
091400
091400
091400
091400
091400
091400
091400
091400
091400
091400
091400
091400
091400
091400
091400
091400
091400
091400
091400
091400
091400
091400
091400
091400
091400
091400
091400
091400
091400
091400
091400
091400
091400
091400
091400
091400
091400
091400
091400
091400
091400
091400
091400
091400
091400
091400
091400
091400
091400
091400
091400
091400
091400
091400
091400
091400
091400
0914000
091400
091400
091400
091400
091400
091400
091400
091400
0914000
091400
091400
091400
091400
091400
091400
091400
091400
0914000
091400
091400
091400
091400
091400
091400
091400
091400
0914000
091400
091400
091400
091400
091400
091400
091400
091400
091400
091400
091400
091400
091400
091400
091400
091400
091400
091400
091400
091400
091400
091400
091400
091400
091400
091400
091400
091400
091400
091400
091400
091400
091400
091400
091400
091400
091400
091400
091400
091400
091400
091400
091400
091400
091400
091400
091400
091400
091400
091400
091400
091400
091400
091400
091400
091400
091400
091400
091400
091400
091400
091400
091400
091400
091400
091400
091400
091400
091400
091400
091400
091400
09140
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        RESULT
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             Q9K9P8
                                                                                                                                                                                                                             g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      QQ
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  q
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     pp
                                                                                                                                                   ò
                                                                                                                                                                                                                                                                                                           ŏ
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        δ
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                δλ
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            δ
```

```
58 IIGHGIPLYRFSSSIVPLATHP-DVMMDFVTPFQKEFREIGELVKTHQLRTSFHPNQFTL 116
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          117 FTSPKESVTKNAVTDMAYHYRMLEAMGIADRSVINIHIGGAYGNKDTATAQFHQNIKQLP 176
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             119 LNSEAKEITRRSLQTLLYHYKLLKGMEIDPRHRCVLHVGGKKKGVEAGLEQFIENTASIP 178
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   177 QEIKERMTLENDDKTYTTEETLQVCEQEDVPFVFDFHHFYANPDDH-----ADLNVALP 230
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          231 RMIKTWERIGLOPKVHLSSPKSEOAIRSHADYVDANFLLERFROWGT----NIDFMIEAK 286
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              232 RIVATWEDSPLPVKIHLSSPLSGEDDPRHHDYINADRFIAFLHEIGADAVDHLHVMIEAK 291
                                                                                                                                                                                                                                                                                               halodurans and genomic sequence comparison with Bacillus subtilis.";
Nucleic Acids Res. 28:4317-4331(2000).
EMBL; AP001516; BAB06316.1; -.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                       Gaps
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         1 MIFREGFVSNAMSLWDASPAKTLTFARYSKLSKTE---RKEALLTVTKANLRNTMRTLHY 57
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             1 MRIQFGYVAMSMELANASPSKTMTATQFEKIEDHEAGLRK--LERIAKTNLHNCLRLLKH 58
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              Takao M., Yonemasu R., Yamamoto K., Yasui A.; "Characterization of a UV endonuclease gene from the fission yeast Schizosaccharomyces pombe and its bacterial homolog.";
                                                                                                                                                                                    STRAIN=C-125 / JCM 9153;
MEDLINBE-2515282; PubMed=11058132;
Takami H., Nakasone X., Takaki Y., Maeno G., Sasaki R., Masui N.,
Fuji F., Hirama C., Nakamuza Y., Ogasawara N., Kuhara S.,
                                                                                                                                                                                                                                                                               "Complete genome sequence of the alkaliphilic bacterium Bacillus
                                                                                                                                                                                                                                                                                                                                                                                                                                  DB 2; Length 322;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                     133; Indels
                                                                                                                                                                                                                                                                                                                                                                          322 AA; 36900 MW; 4B0DC9291CD1E7B4 CRC64;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               Eukaryota, Fungi, Ascomycota, Schizosaccharomycetes, Schizosaccharomycetales; Schizosaccharomycetaceae;
01-OCT-2000 (TrEMBLrel. 15, Last sequence update) 01-OCT-2000 (TrEMBLrel. 15, Last annotation update)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  01-NOV-1996 (TrEMBLrel. 01, Created)
01-NOV-1996 (TrEMBLrel. 01, Last sequence update)
01-NOV-1998 (TrEMBLrel. 08, Last annotation update)
                                                                                           Bacteria; Firmicutes; Bacillus/Clostridium group; Bacillus/Staphylococcus group; Bacillus.
                                                                                                                                                                                                                                                                                                                                                                                                                                29.2%; Score 484.5; DB 2 32.2%; Pred. No. 6.7e-32;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  599 AA.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                   67; Mismatches
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               Schizosaccharomyces pombe (Fission yeast)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   287 QKDKALLRLMDELSSIRGVKRIGGGALQW 315
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     292 KKDLALFQLMKDLAEYDEITVVSKSAVEF 320
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  PRT;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                MEDLINE=96188860; PubMed=8614629;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                     Conservative
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  PRELIMINARY;
                                                                         Bacillus halodurans.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       Schizosaccharomyces.
                                                                                                                                                                                                                                                                                                                                                                                                                                                   Local Similarity
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              SEQUENCE FROM N.A.
                                                                                                                                                                                                                                                                                                                                                        Complete proteome.
                                                                                                                                                                    SEQUENCE FROM N.A.
                                                                                                                                  NCBI_TaxID=86665;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           UV-ENDONUCLEASE.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      NCBI_TaxID=4896;
                 01.-OCT-2000 (TrE
BH2597 PROTEIN.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  Q10988; P87339;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               UVDE OR UVE1+.
                                                                                                                                                                                                                                                               Horikoshi K.;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                STRAIN-SP972;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                     106;
                                                                                                                                                                                                                                                                                                                                                                            SEQUENCE
                                                                                                                                                                                                                                                                                                                                                                                                                                    Query Match
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              01.0988
                                                         BH2597
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           RESULT 4
                                                                                                                                                                                                                                                                                                                                                                                                                                                                       Matches
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                010988
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              qq
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  qq
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        pp
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          Qy
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             Dp
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          Öλ
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   ŏ
 QΥ
```

9

m

```
7;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              "A eukaryotic gene encoding an endonuclease that specifically repairs DNA damaged by ultraviolet light."; EMBO J. 14:2393-2399(1995).
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   454 IPLVLDWHHHNIVPGTLREGSLDLMPLIPTIRETWIRKGITQKQHYSESADPTAISGMKR 513
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             334 QSHLEEVGKLANKYNHRLTMHPGQYTQIASPREVVVDSAIRDLAYHDEILSRMKLNEQLN 393
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              148 --SVINIHIGGAYGNKDTATAQFHQNIKQLPQEIKERMTLENDDKTYTTEETLQVCEQED 205
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  206 VPFVFDFHHFYANP----DDHADLNVALPRMIKTWERIGLQPKVHLSSPKSEQAI---- 256
                                                                                                                                                                                                                                                                                                                                                                                                                                                                    276 TIQRDGLESVKQLGTQNVLDLIKLVEWNHNFGIHFMRVSSDLFPFASH··AKYGYTLEFA 333
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        90 QKEFREIGELVKTHQLRTSFHPNQFTLFTSPKESVTKNAVTDMAYHYRMLEAMGIADR-- 147
                                                                                                                                                                                                                                                                                                                                                                                               Gaps
                                                                                                                                                                                                                                                                                                                                                                                                                                34 TERKEALLTVTKANLRNTM---RTLHYIIGHGIPLYRFSSSIVPLATHPDVMWDFVTPF- 89
                                                                                                                                                                                        Lyne M., Rajandream M.A., Barrell B.G., Lucas M., Gaillardin C.; Submitted (JUN-1998) to the EMBL/GenBank/DDBJ databases. EMBLE, D78571; BAA11415.1: ... EMBLE, U78487; AAC49644.1; -.. EMBLE, D7823859; CAA19577.1; -.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            Yajima H., Takao M., Yasuhira S., Zhao J.H., Ishii C., Inoue H.,
                                                                                                                                                                                                                                                                                                                                                                                             25;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    Eukaryota; Fungi; Ascomycota; Pezizomycotina; Sordariomycetes; Sordariales; Sordariaceae; Neurospora.
                                                                                                                                                                                                                                                                                                                                                             DB 3; Length 599;
                                           MEDIINE-97175806; PubMed-9023111;
Davey S., Nass M.L., Ferrer J.V., Sidik K., Eisenberger A., Antchell D.L., Freyer G.A., "The fission past UVDR DNA repair pathway is inducible.";
Nucleic Acids Res. 25:1002-1008(1997).
                                                                                                                                                                                                                                                                                                                                                         21.6%; Score 359.5; DB 3; Length 29.7%; Pred. No. 2.8e-21; tive 60; Mismatches 111; Indels
                                                                                                                                                                                                                                                                                                      599 AA; 68816 MW; C09ED7219D5895BE CRC64;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         656 AA; 74413 MW; 4B231E7A8735C86F CRC64
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          01-NOV-1996 (TrEMBLrel. 01, Created)
01-MAY-1999 (TrEMBLrel. 10, Last sequence update)
01-MAY-1999 (TrEMBLrel. 10, Last annotation update)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                257 RSHADYVDANFLLERFRQWGTNIDFMIEAKQKDKALLRL 295
Nucleic Acids Res. 24:1267-1271(1996).
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              STRAIN=74-OR23-1A;
MEDLINE=95292980; PubMed=7774597;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          PRT;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       EMBL; D11392; CAB20113.1;
                                                                                                                                                                                                                                                                                                                                                         Query Match
Best Local Similarity 29.7<sup>†</sup>
Matches 83; Conservative
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          PRELIMINARY;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     Neurospora crassa.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             SEQUENCE FROM N.A.
                                                                                                                                                            SEQUENCE FROM N.A.
                                 SEQUENCE FROM N.A.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        NCBI_TaxID=5141;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               UV-ENDONUCLEASE
                                                                                                                                                                               STRAIN=972H-;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          Endonuclease
SEQUENCE 65
                                                                                                                                                                                                                                                                                         Endonuclease
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     Yasui A.;
                                                                                                                                                                                                                                                                                                          SEQUENCE
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          001408
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          셤
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             ö
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                ò
```

```
11;
                8
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            Kitayama S., Kikuchi M., Funayama T., Narumi I., Watanabe H.; "Cloning of structural gene of an alternative incision enzyme for DNA demans in Definionary radiodirans ":
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             81 VMWD-----FVTPFQKEFREIGELVKTHQLRTSFHPNQFTLFTSPKESVTKNAVTDMAYH 135
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                82 LAGDDTGAAVLTHLAPQLLEAGHAFTDAGVRLLMHPEQFIVLNSDRPEVRESSVRAMSAH 141
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   136 YRMLEAMGIADRSVINIHIGGAYGNKDTATAQFHQNIKQLPQEIKERMTLENDDKTYTTE 195
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     196 ETLQVCEQEDVPFVFDFHHFYANP--DDHADLNVA--LPRMIKTWERIGLQPK----VHL 247
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      199 ELLPICEATGTPLVFDAHHHVVHDKLPDQEDPSVREWVLRARATW-----QPPEWQVVHL 253
                                                                                                                                                                                                                                                                                   277 ANARDIVKMLCWNEKYGIRFIRLSSEMFPFASHP-VHGYKLAPFASEVLAEAGRVAAELG 335
                                                                                                                                           Gaps
                                                   46 ANLENTMRTLHYIIGHGIPLYRFSSSIVPLATHPDVMWDFVTPFQKE-FREIGELVKTHQ 104
                                                                                                                       105 LRTSFHPNQFTLFTSPKESVTKNAVTDMAYHYRMLEAMGIADR----SVINIHIGGAYGN 160
                                                                                                                                                                                           161 KDTATAQFHQNIKQLPQEIKERMTLENDDKTYTTEFTLQVCEQEDVPFVFDFHHFYANPD 220
                                                                                                                                                                                                                                                               221 D----HADLNVALP----RMIKTWERIGLQPKVHLSSPKSEQAI-----RSHADYVDANF 267
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           21 KTLTFARYSKLSKTERKEALLTVTKANLRNTMRTLHYIIGHGIPLYRFSSSIVPLATHPD 80
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             Bacteria; Thermus/Deinococcus group; Deinococcales; Deinococcus
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 248 SSPKSEQAIRSHADYVDANFLLERFRQWGTNIDFM-IEAKQKDKAL--LRLM 296
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     Length 305;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        Indels
                  Indels
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        damage in Deinococcus radiodurans.";
Submitted (OCT-1999) to the EMBL/GenBank/DDBJ databases
EMBL; AB033747; BAA85759.1; -
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                305 AA; 33592 MW; B94D333243E2FEA4 CRC64;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           01-MAY-2000 (TrEMBLrel. 13, Created)
01-MAY-2000 (TrEMBLrel. 13, Last sequence update)
01-MAY-2000 (TrEMBLrel. 13, Last annotation update)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   17.7%; Score 293.5; DB 2; 30.1%; Pred. No. 2.9e-16; tive 51; Mismatches 120;
Best Local Similarity 32.3%; Pred. No. 4.9e-20;
Matches 87; Conservative 51; Mismatches 105;
                                                                                                                                                                                                                                                                                                                                                                                                                                                              305 AA
                                                                                                                                                                                                                                                                                                                                     268 LLERFRQWGTNIDFMIEAKQKDKALLRLM 296
                                                                                                                                                                                                                                                                                                                                                           S11 --MTLPPCPPDMDLMIEAKDKEQAVFELM 537
                                                                                                                                                                                                                                                                                                                                                                                                                                                                PRT;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         Best Local Similarity 30.1%
Matches 88; Conservative
                                                                                                                                                                                                                                                                                                                                                                                                                                                                PRELIMINARY;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       Deinococcus radiodurans.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             SEQUENCE FROM N.A.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    UV-ENDONUCLEASE.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          NCBI_TaxID=1299;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       Endonuclease
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              STRAIN=KR1;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   SEQUENCE
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           Query Match
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                ;6M0S60
                                                                                                                                                                                                                                                                                                                                                                                                                                                                6M0S60
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         UVSCDE
                                                                                                                                                                                                                                                                                                                                                                                                                               RESULT
09S0M9
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            qq
                                                                                                                                                                                                                                                                                                                                                                                                                                                                  δ
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 qq
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  a
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        δŏ
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        qq
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            õ
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                δ
                                                                                                                                                                                                                                                                                                      qq
                                                                                                                                                                                                                                                                                                                                        ΟŊ
                                                                                                                                                                                                                                                                                                                                                                          g
                                                                                                                                                            qq
                                                                                                                                                                                                ò
                                                                                                                                                                                                                                qq
                                                                                                                                                                                                                                                                    δ
                                                                                                                            δ
```

STRAIN-972H-;

```
Qy
                                                                                                                                                                                                                                                                      q
                                                                                                                                                                                                                                                                                                   δy
                                                                                                                                                                                                                                                                                                                            qq
                                                                                                                                                                                                                                                                                                                                                           οy
                                                                                                                                                                                                                                                                                                                                                                                                                                               qq
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        q
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    Ó
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         11;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         46 RTVTLSRYRALSPAEREAKLLDLYSSNIKTLRGAADYCAAHDIRLYRLSSSLFPML---D 102
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   81 VMWD-----FVTPPQKEFREIGELVKTHQLRTSFHPNQFTLFTSPKESVTKNAVTDMAYH 135
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 103 LAGDDTGAAVLTHLAPQLLEAGHAFTDAGVRLLMHPEQFIVLNSDRPEVRESSVRAMSAH 162
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       163 ARVMDGLGLA-RTPWNLLL--LHGGKGGRGAELAALIPDLPDPVRLRLGLENDERAYSPA 219
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             136 YRMLEAMGIADRSVINIHIGGAYGNKDTATAQFHQNIKQLPQEIKERMTLENDDKTYTTE 195
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     196 ETLQVCEQEDVPFVFDFHHFYANP--DDHADLNVA--LPRMIKTWERIGLQPK----VHL 247
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               220 ELLPICEATGTPLVFDAHHHVVHDKLPDQEDPSVREWVLRARATW-----QPPEWQVVHL 274
                                                                                                                                                                                                                 MEDLINE=20036896; PubMed=10567266; White O., Eisen J.A., Heidelberg J.F., Hickey E.K., Peterson J.D., Dodson R.J., Haft D.H., Gwinn M.L., Nelson W.C., Richardson D.L., Moffat K.S., Qin H., Jiang L., Pamphile W., Crosby M., Shen M., Vamathevan J.J., Lam P., McDonald L., Utterback T., Zalewski C., Makarova K.S., Aravind L., Daly M.J., Minton K.W., Fleischmann R.D., Ketchum K.A., Nelson K.E., Salzberg S., Smith H.O., Venter J.C.,
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             21 KTLTFARYSKLSKTERKEALLTVTKANLRNTMRTLHYIIGHGIPLYRFSSSIVPLATHPD 80
                                                                                                                                                 Thermus/Deinococcus group; Deinococcales; Deinococcus.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      33;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             248 SSPKSEQAIRSHADYVDANFLLERFRQWGTNIDFM-IEAKQKDKAL--LRLM 296
                                                                                                                                                                                                                                                                                                                                                                                                                                                                   17.7%; Score 293.5; DB 2; Length 326; 30.1%; Pred. No. 3.2e-16; tive 51; Mismatches 120; Indels 33
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    775 SNGIEGPQDRRH-----SHLIADFPSAYADVPWIEVEAKGKEEALAALRLM 320
                                                                                                                                                                                                                                                                                                                                      "Genome sequence of the radioresistant bacterium Deinococcus
                                                                                                                                                                                                                                                                                                                                                                                                                             C4EA0D0AD2C38988 CRC64;
                                                      01-MAY-2000 (TrEMBLrel. 13, Created)
01-MAY-2000 (TrEMBLrel. 13, Last sequence update)
01-WAR-2001 (TrEMBLrel. 16, Last annotation update)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    Eukaryota, Fungi, Ascomycota, Schizosaccharomycetes, Schizosaccharomycetales, Schizosaccharomycetaceae,
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          01-MAR-2001 (TrEMBLrel. 16, Created)
01-MAR-2001 (TrEMBLrel. 16, Last sequence update)
01-MAR-2001 (TrEMBLrel. 16, Last annotation update)
                                  326 AA
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        Schizosaccharomyces pombe (Fission yeast).
                                  PRT;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   PRT;
                                                                                      -MAR-2001 (TrEMBLrel. 16, Last DAMAGE ENDONUCLEASE, PUTATIVE.
                                                                                                                                                                                                                                                                                                                                                                                            TIGR; DR1819; -. Endonuclease; Complete proteome
                                                                                                                                                                                                                                                                                                                                                                                                                           326 AA; 35693 MW;
                                                                                                                                                                                                                                                                                                                                                                                   EMBL; AE002022; AAF11370.1; -.
                                                                                                                                                                                                                                                                                                                                                  radiodurans R1.";
Science 286:1571-1577(1999).
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  Conservative
                                PRELIMINARY;
                                                                                                                                Deinococcus radiodurans.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   PRELIMINARY;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          PUTATIVE DNA HELICASE.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   Local Similarity
Hes 88; Conserve
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  Schizosaccharomyces
                                                                                                                                                                                          SEQUENCE FROM N.A.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            SEQUENCE FROM N.A.
                                                                                                                                                               NCBI_TaxID=1299;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                NCBI_TaxID=4896;
                                                                                                                                                                                                                                                                                                                     Fraser C.M.;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       SPBC543.03C
                                                                                                                                             Bacteria;
                                                                                                                                                                                                                                                                                                                                                                                                                             SEQUENCE
                                                                                                                                                                                                                                                                                                                                                                                                                                                                        Query Match
                                             09RTE6;
                               Q9RTE6
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   09HGM8;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   O9HGM8
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     æ
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               Matches
RESULT
               O9RTE6
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     RESULT
Q9HGM8
                                             g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     δ
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           δλ
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               q
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   ID
DT
DT
DT
DT
DT
OC
OC
OC
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     ŏ
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             οy
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        qq
```

```
10;
                                                                                                                                                                                                                                                             79; Gaps
                                                                                                                                                                                                                                                                                                                                                                      429 AEDVRPYSFSEFEKLSNEE------470
                                                                                                                                                                                                                                                                                                                                                                                                                             77 THPDVMWDFVTPFQKE--FREIGELVKTHQLR--TSFHPNQ------FTLFTSPK 121
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   471 -----GENPPFEPENTFSMI-----PHRLQQAISYYANSPEGDLPQPNIYLTRYINPP 518
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         122 ESVTKNAVTDMAYHYRMLEAMGIADRSVINIHIGGAYGNKDTATAQFHQNIKQLPQEIKE 181
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      182 RMTLENDDKTYTTEETLQVCEQEDVPFVFDFHHFYANPDDHADLNVALPRMIKTWERIGL 241
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            567 LLNSGLDEQEGEKLLVLHVSEKDPVG---TFTEVLKNP---FGLEDALTEMEKVIKNLID 620
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     242 QPKVHLSSPKSEQAIRSHADYVDANFLLERFRQWGTNI--DFMIEAKQKDKALLR---- 294
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       Okinaka R.T., Cloud K., Hamton O., Hoffmaster A., Hill K.K., Keim P., Koehler T., Lamke G., Kumano S., Mahillon J., Manter D., Martinez Y., Ricke D.O., Syensson R., Jackson P.J.;
                                                                                                                                                                                                                                                                                                                17 ASPAKTLIFARYSKLSKTERKEALLTVTKANLRNTMRTLHYIIGHGIPLYRFSSSIVPLA 76
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             "The sequence and organization of pXO1, the large Bacillus anthracis plasmid harboring the Anthrax toxin genes."; J. Bacteriol. 0:0-0(1999).
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               41; Mismatches 114; Indels 65; Gaps
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         92 EFREIGELVKTHQLRTSFHPNQFTLFTSPKESVTKNAVTDMAYHYRMLEAMGIADRSVIN 151
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  32 SKTERKEALLTVTKANLRNTMRTLHYIIGHGIPLYRFSSSIVPLATHPDVMWDFVTPFQK 91
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            519 KSLLDNCILDLKL,-----IKEKLTVNVPVKPKYSSQETA---FDTGAPISEEQIEE
Wood V., Rajandream M.A., Barrell B.G., Saunders D., Harris D.;
Submitted (AUG-1999) to the EMBL/GenBank/DDBJ databases.
EMBL; AL391604; CAC05245.1; -
                                                                                                                                                                                                    6.6%; Score 109.5; DB 3; Length 695; 21.9%; Pred. No. 1.1;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      6.3%; Score 104.5; DB 2; Length 315; 22.8%; Pred. No. 0.98;
                                                                                                                                                                                    Query Match
Best Local Similarity 21.9%; Pred. No. 1.1;
Matches 68; Conservative 50; Mismatches 114; Indels
                                                                                                               695 AA; 79871 MW; 2FA77B1FD1401CAA CRC64;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 315 AA; 37322 MW; 436CCE09A4A46118 CRC64;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          Last sequence update)
Last annotation update)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         Bacteria; Firmicutes, Bacillus/Clostridium group; Bacillus/Staphylococcus group; Bacillus/
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         315 AA
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   01-NOV-1999 (TrEMBLrel. 12, Created)
01-NOV-1999 (TrEMBLrel. 12, Last seq
01-NOV-1999 (TrEMBLrel. 12, Last anno
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         PRT;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    Plasmid virulence plasmid PX01.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        EMBL; AF065404; AAD32322.1; -.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 22.8%;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               65; Conservative
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      PRELIMINARY;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        295 ----LMDELS 300
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           677 GLDIILHDELT 687
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 Best Local Similarity
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      Bacillus anthracis.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               SEQUENCE FROM N.A.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     NCBI_TaxID=1392;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    STRAIN-STERNE;
                                                                                       Helicase.
                                                                                                                  SEQUENCE
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 SEQUENCE
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      Query Match
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           PX01-18.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              09X2Y9;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   09X2Y9
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                σ
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         Matches
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                RESULT
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           09X2Y9
```

```
CG14967
                                                                                                                                                            097257
                                                                                                                                      RESULT 11
                                                                                                                                                   782760
                                                                                                                                                                                  Qγ
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               15;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          157 FRESSFDSNLSTSSSSPPKD----KKASVSDTVSDSSTVTASMSNMPTISIDLNLDEMHD 212
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      213 IIKSPETPTPTTGLPTQKAEKKASPTAIKNWQAPE-SWDVKAPIKKEEPHAPKIEEVAEN 271
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               ------LVKTHQ------LRTSFHPNQFTLFTSPKESVTKNA 128
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      272 DVAIDNVLEKKRLPVLYGTHQVPHVTNSKDIKSSHIIRVFKEDNTFTTILCPLETTTS-- 329
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   3 FRFGFVSNAMSLWDASPAKTLTFARYSKLSKTERKEALLTVTKANLRN-----TMRTLHY 57
                                      197 ILQVCEQEDVPF--VFDFHH----FYANPDDHADLNVALPRMIKTWERIGLQPKVHLSS 249
                                                                                         216 PVRPQQHTNNPLFVAFDFNRCTFRWVYENDAPKALSEVAIQKMIR-----LEV 263
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               113; Gaps
125 -----TLLKQTVLST-----1MDRNACI 155
                        152 IHIGGAYG-----NKDTATAQFHQNIKQLPQEIK----ERMTLENDDKTY----TTEE 196
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   58 IIGH-----GIPLYRFSSSIVPLATH----PDVMMDFVTPFOKE-----FREIGE-
                                                                                                                                                                                                                                                                                          Eukaryota; Fungi; Ascomycota; Saccharomycotina; Saccharomycetes; Saccharomycetales; Candida.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        6.3%; Score 104.5; DB 3; Length 1690; 20.4%; Pred. No. 9.7;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               37; Mismatches 107; Indels
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            GUANYLATE_CYCLASES_2; 1.
A; 189400 MW; 8AB32D9B8C747AB0 CRC64;
                                                                                                                                                                                                                                                                                                                                                                          Submitted (NOV-1999) to the EMBL/GenBank/DDBJ databases.
EMBL: AB034965; BAA93553.1; -.
                                                                                                                  250 PKSEQAIRSHADYVDANFLLERFRQWGTNIDFMIEAKQKDKALLR 294
                                                                                                                                      264 KRAELDRRISAQQMRNTFILRLIRQGVTEKDLVSMMGFKTKISLK 308
                                                                                                                                                                                                                        01-OCT-2000 (TrEMBLrel. 15, Created)
01-OCT-2000 (TrEMBLrel. 15, Last sequence update)
01-JUN-2001 (TrEMBLrel. 17, Last annotation update)
                                                                                                                                                                                                  PRT; 1690 AA.
                                                                                                                                                                                                                                                                                                                                                                                                                       InterPro; IPR003592; LRR_out.
InterPro; IPR003592; LRR_typ.
InterPro; IPR003599; PP2C_catalytic.
InterPro; IPR001932; PP2C_domain.
InterPro; IPR001938; PP2C_sig.
InterPro; IPR001159; RA.
Pfam; PF00211; quanylate_cyc; 1.
Pfam; PF00481; PP2C; 2.
                                                                                                                                                                                                                                                                                                                                                                                                   InterPro; IPR001054; Guanylt_cyclase
                                                                                                                                                                                                                                                                                                                                                        Ono N., Sudoh M.; "Candida albicans CYR1 gene.";
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              PRINTS; PR00019; LEURICHRPT.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           SMART; SM00332; PP2Cc; 1.
SMART; SM00331; PP2C_SIG; 1.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         SMART; SM00044; CYCc; 1.
SMART; SM00370; LRR; 7.
SMART; SM00369; LRR_TYP; 2.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    20.48;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          Query Match 6.3%
Best Local Similarity 20.4%
Matches 66; Conservative
                                                                                                                                                                                                                                                                                   Candida albicans (Yeast).
                                                                                                                                                                                                   PRELIMINARY;
                                                                                                                                                                                                                                                                                                                                                                                                               InterPro; IPR001611;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         1690 AA;
                                                                                                                                                                                                                                                                                                                                           SEQUENCE FROM N.A.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              PROSITE; PS50125;
                                                                                                                                                                                                                                                                                                                    NCBI_TaxID=5476;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         SEOUENCE
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   SMART;
                                                                                                                                                                                                   09P977
                                                                                                                                                                                                               09P977
                                                                                                                                                                            RESULT 10
Q9P977
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           q
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   οy
                                                                                                                                                                                                   ð
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     δ
                                                                                                                                          q
                                                                                             qq
                                                                                                                    δ
                                                                       ò
```

```
RC STRAIN=BERKELEY;

RA Adams M.D., Celniker S.E., Holt R.A., Evans C.A., Gocayue J.D.,

RA Adamstides P.G., Schrer S.E., Holt R.A., Evans C.A., Gocayue J.D.,

RA Adamstides P.G., Schrer S.E., Holt R.A., Evans C.A., Galle R.F.,

RA George R.A., Lewis S.E., Richards S., Ashburner M., Henderson S.M.,

RA George R.A., Lewis S.E., Richards S., Ashburner M., Pfeiffer B.D.,

RA Brandon R.C., Rogers Y.-H.C., Blazel R.G., Champe M., Pfeiffer B.D.,

RA Brandon R.C., Robylan A., Marger E.G., Helf G., Nelson C.R., Malklos G.L.G.,

RA Ballew R.M., Basu A., Baxendale J., Bayeaktaroqlu L., Beasley E.M.,

RA Ballew R.M., Basu A., Baxendale J., Bayeaktaroqlu L., Beasley E.M.,

RA Ballew R.M., Basu M., Baxendale J., Brokstein P., Brottief P.,

RA Bertis K.C., Busam D.A., Bulck H., Cadleu E., Center A., Chandra I.,

RA Bertis K.C., Busam D.A., Buller H., Cadleu E., Center A., Chandra I.,

RA Glodson K., Doubp L.E., Downes M. Duyan R.C., Perrice S., Plaksvor B.C.,

RA Glodson K., Doubp L.E., Downes M., Duyan R.C., Perrica S., Fleischman W.,

RA Glodek A., Gong F., Gorriellu H., Golbart W.M., Classer K.,

RA Harris N.L., Harvey D., Heiman T.J., Herrandez J.R., Houck J.,

RA Harris N.L., Harvey D., Heiman T.J., Wei M.-H., Ibegwam C.,

RA Jalai M., Kalush F., Kappen G.H., Ka Z., Kulp D., Lai Z.,

RA Jalai M., Malush F., Kappen G.H., Walvitz S., Kulp D., Lai Z.,

RA Liu X., Mattei B., WcIntosh T.C., McLeod M.P., Neilm H., Nelson D.L.,

Mount S.M., Moy M., Murphy B., Murphy L., Wurny D.M., Nelson D.L.,

RA Reibert K., Remington K., Saunders R., Venis D., Puri V., Kapper E., Spradling A.C., Pan S., Pollard J., Puri V., Wesser M.,

RA Spier E., Spradling A.C., Stapleton M., Stuppski M.P., Smith H.O.,

RA Spier E., Spradling A.C., Stapleton M., Stuppski M.P., Smith H.O.,

RA Sher B., Jane S., Woolge T., Worley K.C., Wu D., Yang S., Yang O., Yang S., Jang S., Pan R., Pollar R., Pan S., Pollar W., Shill M., Pan S., 
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               Eukaryota; Metazoa; Arthropoda: Tracheata; Hexapoda; Insecta; Pterygota; Neoptera: Endopterygota; Diptera; Brachycera; Muscomorpha; Ephydroidea; Drosophilidae; Drosophila.
                                                                                                                                                                184 TLENDDKTYTTEETLQVCEQEDVPFVFDFHHFYANPDDHADLNVALPRMIKTWERIGLQP 243
                                                                                                                                                                                                            129 VIDMAYHYRMLEAMGIADR----SVINIHIGGAYGNKDTATAQFHQNIKQLPQEIKERM 183
                                                                                  330 -----ELLAIVQKKFFLESTTNFQLSVCIGNCVKVLEDFEK----PLKIQMGL 373
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         2272 AA; 259619 MW; 4102D7EC483A3298 CRC64;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         Last sequence update)
Last annotation update)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               PRT; 2272 AA.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           Created)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              Drosophila melanogaster (Fruit fly).
                                                                                                                                                                                                                                                                                                                                              244 KVHLSSPKSEQAIRSHADYVDAN 266
                                                                                                                                                                                                                                                                                                                                                                                                                   409 LTH----DEEVLLSR-NYVDVN 425
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              EMBL, AE003477, AE4474001; -
EMBL, AE003477, AE447401; -
FlyBase, FBgn0035420; CG14967.
SEQUENCE 2272 AA; 259619 MW
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  01-MAY-2000 (TrEMBLrel. 13, 01-MAY-2000 (TrEMBLrel. 13, 01-JUN-2000 (TrEMBLrel. 14,
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        PRELIMINARY;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      SECUENCE FROM N.A.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                CG14967 PROTEIN.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   NCBI_TaxID=7227;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             STRAIN=BERKELEY;
```

```
15;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            RA Mallet L., Remault G., Jacquet M.;

RA Mallet L., Remault G., Jacquet M.;

RT "Functional cloning of the adenylate cyclase gene of Candida albicans
RT in Saccharomyces cerevisiae within a genomic fragment containing five
The agenes, including homologues of CHS6 and SAP185.";

Veast 16:959-966(2000)

REMEL, AJS0877; CAB60230.2;

RINTEPPRO; IPR001054; Guanylt_cyclase.

RINTEPPRO; IPR001054; Guanylt_cyclase.

RINTEPPRO; IPR001591; LRR typ.

RINTEPPRO; IPR001592; LRR typ.

RINTEPPRO; IPR001593; PP2C_catalytic.

BR InterPro; IPR001593; PP2C_catalytic.

BR InterPro; IPR00159; RA.

BR InterPro; IPR00159; RA.

BR Ffam; PF0041; guanylate_cyc; 1.

BR SMART; SM00104; CYCC; 1.

BR SMART; SM00170; LRR; 7.

BR SMART; SM00170; LRR; 7.

BR SMART; SM00130; LRR; 7.

BR SMART; SM00130; LRR; 7.

BR SMART; SM00130; LRR; 7.

BR SMART; SM00131; PP2C_SIG; 1.

BR SMART; SM00131; PP2C_SIG; 1.

BR SMART; SM00131; PP2C_SIG; 1.

BR SMART; SM00149; RA; 1.

BR SMART; SM00149; RA; 1.

BR SMART; SM00149; RA; 1.

BR SMART; SM00140; RA; 1.

BR SMART; SM0140; RA; 1.

BR SMART; SM00140; RA; 1.

BR SMART; SM00140; RA; 
                                                                                                                                                                                                                                                1472 P--THKLVVYDLKGAWTKSNRDVAFALFDSFWKSOKLKNNLSTEAVKSYRKEGPNSAVLK 1529
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       : :| : :| : :| | | | : ::| | : ::| | ::| | ::| | ::| | ::| | ::| | ::| | ::| | ::| | ::| | ::| | ::| | ::| | ::| | ::| | ::| | ::| | ::| | ::| | ::| | ::| | ::| | ::| | ::| | ::| | ::| | ::| | ::| | ::| | ::| | ::| | ::| | ::| | ::| | ::| | ::| | ::| | ::| | ::| | ::| | ::| | ::| | ::| | ::| | ::| | ::| | ::| | ::| | ::| | ::| | ::| | ::| | ::| | ::| | ::| | ::| | ::| | ::| | ::| | ::| | ::| | ::| | ::| | ::| | ::| | ::| | ::| | ::| | ::| | ::| | ::| | ::| | ::| | ::| | ::| | ::| | ::| | ::| | ::| | ::| | ::| | ::| | ::| | ::| | ::| | ::| | ::| | ::| | ::| | ::| | ::| | ::| | ::| | ::| | ::| | ::| | ::| | ::| | ::| | ::| | ::| | ::| | ::| | ::| | ::| | ::| | ::| | ::| | ::| | ::| | ::| | ::| | ::| | ::| | ::| | ::| | ::| | ::| | ::| | ::| | ::| | ::| | ::| | ::| | ::| | ::| | ::| | ::| | ::| | ::| | ::| | ::| | ::| | ::| | ::| | ::| | ::| | ::| | ::| | ::| | ::| | ::| | ::| | ::| | ::| | ::| | ::| | ::| | ::| | ::| | ::| | ::| | ::| | ::| | ::| | ::| | ::| | ::| | ::| | ::| | ::| | ::| | ::| | ::| | ::| | ::| | ::| | ::| | ::| | ::| | ::| | ::| | ::| | ::| | ::| | ::| | ::| | ::| | ::| | ::| | ::| | ::| | ::| | ::| | ::| | ::| | ::| | ::| | ::| | ::| | ::| | ::| | ::| | ::| | ::| | ::| | ::| | ::| | ::| | ::| | ::| | ::| | ::| | ::| | ::| | ::| | ::| | ::| | ::| | ::| | ::| | ::| | ::| | ::| | ::| | ::| | ::| | ::| | ::| | ::| | ::| | ::| | ::| | ::| | ::| | ::| | ::| | ::| | ::| | ::| | ::| | ::| | ::| | ::| | ::| | ::| | ::| | ::| | ::| | ::| | ::| | ::| | ::| | ::| | ::| | ::| | ::| | ::| | ::| | ::| | ::| | ::| | ::| | ::| | ::| | ::| | ::| | ::| | ::| | ::| | ::| | ::| | ::| | ::| | ::| | ::| | ::| | ::| | ::| | ::| | ::| | ::| | ::| | ::| | ::| | ::| | ::| | ::| | ::| | ::| | ::| | ::| | ::| | ::| | ::| | ::| | ::| | ::| | ::| | ::| | ::| | ::| | ::| | ::| | ::| | ::| | ::| | ::| | ::| | ::| | ::| | ::| | ::| | ::| | ::| | ::| | ::| | ::| | ::| | ::| | ::| | ::| | ::| | ::| | ::| | ::| | ::| | ::| | ::| | ::| | ::| | ::| | ::| | ::| | ::| | ::| | ::| | ::| | ::| | ::| | ::| | ::| | ::| | ::| | ::| | ::| |
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      1582 VYSDDHSTQSRELQLQGLQACSAQDI------IHENWSISLVNSQVLLKGCE 1627
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       1628 TSGY---VIISAAKAEILQREHRPVWRERSLISKTTWKGLLECMOYYATVSAGDNNSLLE 1684
                                                                                                                                                                                                                                                                                                                                                                      74 PLATHPDVMWDFVTPFQKEFREIG-----ELVKTHQLRTSFHPNQFTLF--TSPKESVTK 126
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       127 NAVTDMAYHYRMLEAMGIADRSVINIHIGGAYGNKDTATAQFHQN--IKQLPQEIKERMT 184
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        185 LENDDKTYTTEET----LQVCEQEDVPFVFDFHHFYANPDDHADLNVALPR---MIKTWE 237
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              238 RIGLQPKVHLSSPKSEQAIRSH-------ADYVDANFLLE 270
                                                                                                                 92; Gaps
                                                                                                                                                                                17 ASPAKTLIFARYSKLSKTE---RKEALLIVIKANLRNIMRTLHYIIGHGIPLYRFSSSIV 73
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    Candida albicans (Yeast).
Eukaryota; Fungi; Ascomycota; Saccharomycotina; Saccharomycetes;
Saccharomycetales; mitosporic Saccharomycetales; Candida.
Query Match 6.3%; Score 104; DB 5; Length 2272; Best Local Similarity 19.8%; Pred. No. 16; Matches 68; Conservative 54; Mismatches 129; Indels 9
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               271 RFRQWGTNIDFMIEAKQKDKALLRLMDELSSIRGVKRIGGGAL 313
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  01-007-2000 (TrEMBLrel. 15, Created)
01-007-2000 (TrEMBLrel. 15, Last sequence update)
01-JUN-2001 (TrEMBLrel. 17, Last annotation update)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  PRT; 1690 AA
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       STRAIN=ATCC26555;
MEDLINE=20330545; PubMed=10870107;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             PRELIMINARY;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                SEQUENCE FROM N.A.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            ADENYLATE CYCLASE
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        NCBI_TaxID=5476;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        Q9P411
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      RESULT 12
09P411
                                                                                                                                                                                                                                                                                                                                                                           ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            ŏ
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         ò
```

```
6.1%; Score 101.5; DB 3; Length 1690; 20.4%; Pred. No. 17; Live 36; Mismatches 108; Indels 113; Gaps
                                                                                                      157 FRESSEDSNLSTSSSSPPKD----KKVSVSDTVSDSSTVTASMSNMPTISIDLNLDEMHD 212
                                                                                                                                                                272 DVAIDNVLEKKRLPVLYGTHQVPHVTNSKDIKSSHIIRVFKKEDNTFTTILCPLETTTS-- 329
                                                                                                                                                                                                                                                                                    129 VTDMAYHYRMLEAMGIADR----SVINIHIGGAYGNKDTATAQFHQNIKQLPQEIKERM 183
                                                                                                                                                                                                                                                                                                                                                          184 TLENDDKTYTTEETLQVCEQEDVPFVFDFHHFYANPDHADLNVALPRMIKTWERIGLQP 243
                                                                                                                                                                                                                330 ------PLKIQMKFFLESTTNFQLSVCIGNCVKVLEDFEK-----PLKIQMGL 373
                                                                                                                                                                                                                                                                                                                                                                                            374 LLLSG---YTEEDKLRMLGREDLSFVCKF-------VVENIFLRS 408
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          Rocha C.R.C., Leberer E., Dignard D., Harcus D., Thomas D.Y., "Candida albicans adenylyl cyclase signaling growth, hyphal formation
                                                                      3 FREGEVSNAMSLWDASPAKTLPFARYSKLSKTERKEALLTVTKANLRN-----TMRTLHY 57
                                                                                                                                          58 IIGH------GIPLYRFSSSIVPLATH----PDVMMDFVTPFQKE-----FREIGE- 98
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   Eukaryota; Fungl; Ascomycota; Saccharomycotina; Saccharomycetes; Saccharomycetales; Candida.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         GUANYLATE_CYCLASES_2; 1.
A; 189426 MW; 647468CC0E69F4E4 CRC64;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           Submitted (AUG-2000) to the EMBL/GenBank/DDBJ databases.
EMBL, AF295379; AAG18428.1; -.
InterPro; IPR011054; Guanylt_cyclase.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         01-MAR-2001 (TrEMBLrel. 16, Created)
01-MAR-2001 (TrEMBLrel. 16, Last sequence update)
01-JUN-2001 (TrEMBLrel. 17, Last annotation update)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                PRT; 1690 AA.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  Interpro; IPR001611; LRR.
Interpro; IPR003592; LRR_out.
Interpro; IPR003593; LRR_typ.
Interpro; IPR003599; PP2C_domain.
Interpro; IPR001932; PP2C_domain.
Interpro; IPR001932; PP2C_sig.
Interpro; IPR00199; RA.
Piam; PF00211; guanylate_cyc; 1.
Piam; PF00501; guanylate_cyc; 1.
                                                                                                                                                                                                                                                                                                                                                                                                                                    244 KVHLSSPKSEQAIRSHADYVDAN 266
                                                                                                                                                                                                                                                                                                                                                                                                                                                                    409 LTH----DEEVLLSR-NYVDVN 425
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           Pfam: PF00481; PP2C; 2.
PRAINTS: PR00019; LEURTCHRPT.
SMART: SM00340; LER. 12.
SMART: SM00370; LRR: 12.
SMART: SM00332; PP2Cc; 1.
SMART: SM00331; PP2Cc; 1.
SMART: SM00331; PP2Cc; 1.
PR0SITE: PS50125; GUANYLATE_CYC
SEOUENCE: L690 AA; 189426 WM;
Query Match
Best Local Similarity 20.4%
Matches 66; Conservative
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                Candida albicans (Yeast).
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             PRELIMINARY;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       SEQUENCE FROM N.A.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    ADENYLYL CYCLASE.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       NCBI_TaxID=5476;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  and virulence.";
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           Q9HFT8
Q9HFT8;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         RESULT 13
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            O9HFTB
                                                                         δλ
                                                                                                                                            Qγ
                                                                                                                                                                                                                                                  Q
D
                                                                                                                                                                                                                                                                                      ŏ
                                                                                                                                                                                                                                                                                                                        qq
                                                                                                                                                                                                                                                                                                                                                                                              90
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             οy
                                                                                                                                                                                                                                                                                                                                                                                                                                    Ωÿ
```

```
09KB81;
                                                                                                                                                                                                                                                            BH2047
                                                                                                                                                                        Q9KB81
                                                                                                                                             RESULT 15
                                                                                                                                                            Q9KB81
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        Ωp
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        Op
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      ŏ
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  q
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             δλ
                                       qq
                                                                                                 pp
                                                                                                                                                                                       δ
                                                                     οy
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            13;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         Sensen C.W., Klenk H.-P., Singh R.K., Allard G., Chan C.C.-Y.,
Liu Q.Y., Penny S.L., Young F., Schenk M.E., Gaasterland T.,
Doolittle W.F., Ragan M.A., Charlebois R.L.;
"Organizational characteristics and information content of an archaeal
genome: 156 kb of sequence from Sulfolobus solfataricus P2.";
Mol. Microbiol. 22:175-191(1996).
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           59 ETGQVARELGVRLSVHAPYFINLCSEEKEKIEASKQRILDTADRAELMGADAIA----- 112
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  -TLENDDKTYTTEETL 198
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            113 IHI--AFYGKMT-PEECYONVKEGLEEVIDKAREMGIRNVKFGVETMAKETAFGTLDEVI 169
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         199 QVCEQED--VPFVFDFHHFYANPDDHADLNVALPRMIKTWERIGLQPKVHLSSPKSEQAI 256
                                                                              157 FRFSSFDSNLSTSSSSPPKD----KKVSVSDTVSDSSTVTASMSNMPTISIDLNLDEMHD 212
                                                                                                                                                                                                     272 DVAIDNVLEKKRLPVLYGTHQVPHVTNSKDIKSSHIIRVFKEDNTFTTILCPLETTTS-- 329
                                                                                                                                                                                                                                  129 VTDMAYHYRMLEAMGIADR-----SVINIHIGGAYGNKDTATAQFHQNIKQLPQEIKERM 183
                                                                                                                                                                                                                                                            330 ------ELLAIVQKKFFLESTTNFQLSVCIGNCVKVLEDFEK-----PLKIQMGL 373
                                                                                                                                                                                                                                                                                          184 TLENDDKTYTTEETLQVCEQEDVPFVFDFHHFYANPDDHADLNVALPRMIKTWERIGLQP 243
                                                                                                                                                                                                                                                                                                           95 EIGELVKTHQLRTSFH-PNQFTLFTSPKESV--TKNAVTDMAYHYRMLEAMGIADRSVIN 151
                                                                                                                                                                           -LRTSFHPNQFTLFTSPKESVTKNA 128
                                                                                                                                    36; Mismatches 108; Indels 113; Gaps
                                                                                                                 58 IIGH------GIPLYRESSSIVPLATH----PDVMMDFVTPFQKE-----FREIGE- 98
                                                         3 FRFGFVSNAMSLWDASPAKTLTFARYSKLSKTERKEALLTVTKANLRN-----TMRTLHY 57
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      Archaea; Crenarchaeota; Sulfolobales; Sulfolobus
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            54;
 DB 3; Length 1690;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             6.1%; Score 101; DB 1; Length 283; 21.6%; Pred. No. 1.6;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            Indels
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       PROSITE; PS00729; AP_NUCLEASE_F2_1; UNKNOWN_1.
SEQUENCE 283 AA; 32015 MW; DBFEB9C13F0B5975 CRC64;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                Last annotation update)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            87;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 sequence update)
                                                                                                                                                                                                                                                                                                                                                                                                                                                        283 AA
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            Mismatches
6.1%; Score 101.5; 20.4%; Pred. No. 17;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  152 IHIGGAYGNKDTATAQFHQNIKQLPQEIKERM--
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             EMBL; Y08256; CAA69434.1; -...
InterPro; IPR001119; AP_endonulease_2.
Pfam; PF01261; AP_endonulease2; 1.
SMART; SM00518; APZEC; 1.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   Created)
                                                                                                                                                                                                                                                                                                                                                                                                                                                        PRT;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                SEQUENCE FROM N.A.
STRAIN-DSM 1617 / P2;
MEDLINE-97055432; PubMed-8899719;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    Last
                                                                                                                                                                                                                                                                                                                                                    244 KVHLSSPKSEQAIRSHADYVDAN 266
                                                                                                                                                                                                                                                                                                                                                                               409 LTH----DEEVLLSR-NYVDVN 425
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            48;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                01-JUN-2001 (TrEMBLrel. 17,
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   03,
03,
                                                                                                                                                                           --- LVKTHQ-
   Query Match 6.1%
Best Local Similarity 20.4%
Matches 66; Conservative
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             Conservative
                                                                                                                                                                                                                                                                                                                                                                                                                                                        PRELIMINARY;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  (TrEMBLrel.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           Sulfolobus solfataricus.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     (TrEMBLrel.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         Best Local Similarity
Matches 52; Conserv
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    NCBI_TaxID=2287;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   01-MAY-1997
01-MAY-1997
                                                                                                                                                                             . . . . 66
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  Query Match
                                                                                                                                                                                                                                                                                                                                                                                                                                                     P95930
                                                                                                                                                                                                                                                                                                                                                                                                                              RESULT 14
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     QQ
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              q
                                                                                                                                                                                                                                                                                                                                                                                                                                                                       ò
                                                                                                                                                                                                                                                                                                                                                                                g
                                                                                                                                                                             ö
                                                                                                                                                                                                        g
                                                                                                                                                                                                                                     ò
                                                                                                                                                                                                                                                            g
                                                                                                                                                                                                                                                                                             ò
                                                                                                                                                                                                                                                                                                                       g
                                                                                                                                                                                                                                                                                                                                                     ò
                                                                                                                    ò
```

```
181 ERMTLENDDKTYTTEETLQVCEQEDVPFVFDFHHFYANPDDHADLNVALPRMIKTWERIG 240
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        516 QSIKTEIEEAITHEVASNPKS--VVHLSLDYQD----PDRUMAWLTSPDLM1-AKKDDDS 568
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      142 MGIADRSVI------NIHIGGAYGN------KDTATAQFH--QNIKQ---LPQEIK 180
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        484 KMM ------TPNEDS-----TWEKRW 515
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               378 QERLEWLWKVKQAYPYMTLGQKQELFHKLESTFTEWTGEQP-------RRLGG 423
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                424 EIYADRNLLYEECHGPLNNIKIGGSIKHVLKQDVPKWLSICAKHGEQRRKQEQALAQEIF 483
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  241 LQPKVHL-----SSPKSEQAIRSHADYVDANFLLERFRQWGTNIDFMIEAKQKDKA 291
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               332 KLRLLEKKRLILIEIPLSITEPDSLNELRS--WILK------IKDVEHPDLSLW 377
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            84 DEVTPFQKEFREIGELVKTHQLRTSFH--PNQFTLFTSPKESVTKNAVTDMAYHYRMLEA 141
                                      44; Mismatches 101; Indels 104; Gaps
----ANFLLERFRQWGTNIDFMIEAKQKDKALLRLMDELSS 301
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         "Complete genome sequence of the alkaliphilic bacterium Bacillus halodurans and genomic sequence comparison with Bacillus subtilis."; Nucleic Acids Res. 28:4317-4331(2000).
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        30 KLSKTERK-----EALLTVTKANLRNTMRTLHYIIGHGIPLYRFSSSIVPLATHPDV-MW 83
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        Takami H., Nakasone K., Takaki Y., Maeno G., Sasaki R., Masui
Fuji F., Hirama C., Nakamura Y., Ogasawara N., Kuhara S.,
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             6.0%; Score 100; DB 2; Length 823; Best Local Similarity 23.1%; Pred. No. 8.5; Matches 75; Conservation
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          823 AA; 97055 MW; A661C8566FE8EDC7 CRC64;
                                                                                                                                                                                                                                                                                                                                                                                                         01-OCT-2000 (TrEMBLrel. 15, Created)
01-OCT-2000 (TrEMBLrel. 15, Last sequence update)
01-OCT-2000 (TrEMBLrel. 15, Last annotation update)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                Bacteria: Firmicutes; Bacillus/Clostridium group; Bacillus/Staphylococcus group; Bacillus.
NCBL_TaxID=86665;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                Search completed: January 15, 2002, 14:06:28 Job time: 967 sec
                                                                                                                                                                                                                                                                                                                                               PRT;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 MEDLINE=20512582; PubMed=11058132;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               292 LLRLMDELSSIRGVKRIGGGALQW 315
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         569 YQVILGE---IHDTIMVWGWALQF 589
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       SEQUENCE FROM N.A.
STRAIN=C-125 / JCM 9153;
                                                                                                                                                                                                                                                                                                                                                  PRELIMINARY;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                Bacillus halodurans.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   Complete proteome.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   BH2047 PROTEIN.
         257 RSHADYVD-
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               Fuji F., Hira
Horikoshi K.;
                                                                                                                                302 I 302
                                                                                                                                                                                          278 L 278
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                SEQUENCE
```

!

us-09-724-296-38.rai

```
GenCore version 4.5

Copyright (c) 1993 - 2000 Compugen Ltd.

OM protein - protein search, using sw model

Run on: January 15, 2002, 13:51:16; Search time 57.72 Seconds

(without alignments)
226.125 Million cell updates/sec

Title: US-09-724-296-38

Perfect score: 3055
Sequence: 1 MGTTGLESLSLGDRGAAPTV......HKLNKYHDWLWENHEKL.SLS 580

Scoring table: BLOSUM62
Gapop 10.0, Gapext 0.5

Searched: 212252 seqs, 22503292 residues

Total number of hits satisfying chosen parameters: 212252
```

1: /cgn2_6/ptodata/2/iaa/5A_COMB.pep:*
2: /cgn2_6/ptodata/2/iaa/5B_COMB.pep:*
3: /cgn2_6/ptodata/2/iaa/6B_COMB.pep:*
4: /cgn2_6/ptodata/2/iaa/6B_COMB.pep:*
5: /cgn2_6/ptodata/2/iaa/PCTUS_COMB.pep:*
6: /cgn2_6/ptodata/2/iaa/PCTUS_COMB.pep:*
7 /cgn2_6/ptodata/2/iaa/PCTUS_COMB.pep:*
7 /cgn2_6/ptodata/2/iaa/Packfiles1.pep:*
7 /cgn2_6/ptodata/2/iaa/Packfiles1.pep:*

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Minimum DB seq length: 0 Maximum DB seq length: 200000000 Issued_Patents_AA:*

Database

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

ADDRESSEE: Bell Seltzer Park & Gibson STREET: 310 UCB Plaza, 3605 Glenwood Avenue, Po Drawer 31107

CORRESPONDENCE ADDRESS:

Raleigh

STATE:

COMPUTER READABLE FORM:
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.25 (EPO)
CURRENT APPLICATION DATA:

APPLICATION NUMBER: US/08/755,587 FILING DATE: 25-NOV-1996

PRIOR APPLICATION DATA:
APPLICATION NUMBER: GB 9523959.6
FILING DATE: 23.NOV-1995
PRIOR APPLICATION DATA:
APPLICATION NUMBER: GB 9525555.0
FILING DATE: 14-DEC-1995

FILING DATE: 14-DEC-1995
PRIOR APPLICATION DATA:
APPLICATION NUMBER: GB 9617961.9
FILING DATE: 28-AUG-1996
ATTORNEY/AGENT INFORMATION:
NAME: Kenneth D Sibley

4.8%; Score 146.5; DB 3; Length 1579; 21.1%; Pred. No. 5.8e-05;

Query Match Best Local Similarity

REFERENCE/DOCKET NUMBER: 5405-135 INFORMATION FOR SEQ ID NO: 184: SEQUENCE CHARACTERISTICS:

: 1579 amino acids amino acid

linear

; LENGTH: 15 ; TYPE: amin ; TOPOLOGY: US-08-755-587-184

REGISTRATION NUMBER: 31,665

g to the	n US/08755587 Phillip A Richard F Alan Michael R Materials and methods relating to the	n US/08755 Phillip A Richard F Alan Michael F	pplicatio 997 ATION: AUTEAB!, Wooster, Ashworth, Stratton, ENTION:	RESULT 1 US-08-755-587-184 : Sequence 184, Applica : Patent No. 604597 : GENERAL INFORMATION: APPLICANT: Futrea APPLICANT: Ashwort APPLICANT: Ashwort APPLICANT: Ashwort APPLICANT: Ashwort APPLICANT: Ashwort APPLICANT: ASHWORT TITLE OF INVENTION	RESULTUS-08-7 1 Seque 2 Pater 3 Pater 4 Pater 3 AR
	ALIGNMENTS				
1,					
Sequence 4, Appli	-353		4 . 4.	103	44
4	-487 -328 -353		4 4 4	103.5 103 103	44 45
+	-08-568 -08-487 -08-328 -08-353		4 4 4 4	103.5 103.5 103 103	444 2444 3448
70	US-08-993-228-19 US-08-993-228-19 US-08-568-459A-4 US-08-328-254-6 US-08-353-700-1		4 4 4 4 4 4	103.5 103.5 103.5 103.5 103	4 4 4 4 4 4 5 5 5 5 5 5 5 5 5 5 5 5 5 5
2,	US-08-484-105-2 US-08-484-106-2 US-08-993-228-19 US-08-568-459A-4 US-08-328-24-6 US-08-353-700-1		444444	103.5 103.5 103.5 103.5 103.5 103	39 4 4 4 4 4 4 4 4 4 4 4 4 4 4 3 4 4 4 4 5 4 4 5 5 4 6 5 6 6 6 6 6 6 6 6
, , , , , , , , , , , , , , , , , , ,	US-08-596-291-3 US-09-100-804-3 US-08-484-105-2 US-08-993-228-19 US-08-993-228-19 US-08-487-4 US-08-328-284-6 US-08-328-254-6 US-08-353-700-1		<u>ਰਿ</u> ਰਾਚਦਾਰਾਚਰ	104 103.5 103.5 103.5 103.5 103.5 103.5	7 8 8 8 8 8 8 8 8 8 8 8 8 8 8 8 8 8 8 8
72,33	US-08-986-106-2 US-08-596-291-3 US-09-100-804-3 US-08-484-105-2 US-08-993-228-19 US-08-568-459A-4 US-08-328-24 US-08-358-45 US-08-353-700-1		444444444	104 104 103.5 103.5 103.5 103.5 103.5 103.5	00000000000000000000000000000000000000
	US-08-09-753D-2 US-08-09-09-753D-2 US-08-986-106-2 US-09-100-804-3 US-09-100-804-3 US-08-484-106-2 US-08-99-3-28-19 US-08-568-459A-4 US-08-358-24-6 US-08-353-700-1		- 4 4 4 4 4 4 4 4 4 4	104 104 104 103 : 5 103 : 5 103 : 5 103 : 5 103 : 5	4 8 8 8 8 8 8 8 8 8 8 8 8 8 8 8 8 8 8 8
	US-08-825-886-16 US-08-603-753D-2 US-09-099-753-2 US-08-986-106-2 US-08-100-804-3 US-08-100-804-3 US-08-484-106-2 US-08-993-228-19 US-08-568-459A-4 US-08-38-24-6 US-08-353-700-1		<	100 100 100 100 100 100 103 103 103 103	8 8 8 8 8 8 8 8 8 8 8 8 8 8 8 8 8 8 8
	US-08-425-061-16 US-08-6825-886-16 US-08-693-753D-2 US-09-993-753-2 US-08-986-106-2 US-08-986-106-2 US-08-484-105-2 US-08-484-105-2 US-08-484-105-2 US-08-484-105-2 US-08-484-105-2 US-08-568-459A-4 US-08-568-459A-4 US-08-328-24-6 US-08-338-254-6 US-08-338-254-6 US-08-338-254-6		चर ः च्व्वच्चच्चच्चच	100 100 100 100 100 100 103 103 103 103	2 8 8 8 8 8 8 8 8 8 8 8 8 8 8 8 8 8 8 8
	US-08-755-587-186 US-08-425-061-16 US-08-825-886-16 US-08-603-753D-2 US-09-99753-2 US-08-996-106-2 US-08-996-106-2 US-08-100-804-3 US-08-484-105-2 US-08-993-228-19 US-08-993-228-19 US-08-568-459A-4 US-08-328-254-6 US-08-338-254-6 US-08-338-254-6 US-08-338-254-6		विवर प्रविवव्यव्यव्यव्यव्यव्यव्यव्यव्यव्यव्यव्यव्य	001 001 001 001 001 001 001 001 001 001	
86, 10, 10, 10, 10, 10, 10, 10, 10, 10, 10	953-992A-8 7755-587-18 825-886-16 825-886-16 809-753D-2 996-106-2 996-106-2 996-106-2 998-106-2 999-3-228-19 568-459A-4 4487-828-19 558-459A-4			001 001 1001 1001 1001 1001 1001 1001	0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0

```
19;
                                                                                               306 KSDISLNIDKIPEKNNDYMNKWAG-----LLGPISNHSFGGSFRTASNKEIKLSEHNIKK 360
                                                                                                                                                    96 QRLFGKTAGRFDVYFISPQGLKFRSKSSLANYLHKNGETSLKPEDFDFTVLSKRGIKSRY 155
                                                                                                                                                                                                                                                   156 KDCSMAALTSHL--QNQSNNSNWNLRTRSKCKKDVFMPPSSSSELQESRGLSNFTSTHLL 213
                                                                                                                                                                                                                                                                                                                                                    214 LKEDEGVDDVNFRKVRK-------PKGKVTILK-------239
                                                                                                                                                                                                                                                                                                                                                                                                 452 --- EESGSOFEFTQFRKPSYILQKSTFEVPENQMTILKTTSEECRDADLHVIMNAPSIQV 508
                                                                                                                                                                                                                                                                                                                                                                                                                                                    240 -----GIPIKK-----TKKGCRKSCSGFVQSDSKRESVCNKADAESEPVAQKSQLD- 285
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   509 DSSKQFEGTVEIKRKFAGLLKNDCNKSASGYLTDE-----NEVGFRGFYSAHGTKLNV 561
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   286 RTVCISDAGACGETLSVTSEENSLVKKKERSLSSGSNFCSEQKTSGIINKFCSAKDSEHN 345
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     562 STEALQKAVKLFSDIENISEETS---AEVHPISLSSSKCHDSVVSMFKIENHNDKTSEKN 618
                                                                                                                                                                                                361 SKMFFK-----DIEEQYPTSL---ACVEIVNTLALDNOKKLSKPQIN-TVSAHLQSSVVV 411
                                                                                                                                                                                                                                                                                                  412 SDCKNSHITPQMLFSKQDFNSNHNL------TPSQKAEITE-----LSTIL- 451
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     346 EKYEDTFLESEEI--GTKVEVVERKEHLHTDILKRGSEMDNN--CSPTRK----DFTGEK 397
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       619 NKCQLILQNNIEMTTGTEVEEI------TENYKRNTENEDNKYTAASRNSHNLEFDGSD 671
94; Conservative 67; Mismatches 170; Indels 115; Gaps
                                               36 KEDVAMELERVGEDEEQMMIKRSSECNPLLQEPIASAQFGATAGTECRKSVPCGWERVVK 95
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   APPLICANT: Rommens, Johanna
APPLICANT: Weber, Barbara
TITLE OF INVENTION: Chromosome 13-Linked Breast Cancer
TITLE OF INVENTION: Susceptibility Gene
NUMBER OF SEQUENCES: 124
CORRESPONDENCE ADDRESS:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          Venable, Baetjer, Howard & Civiletti
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        SOFTWARE: Patentin Release #1.0, Version #1.30 CURRENT APPLICATION DATA:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            1201 New York Avenue N.W., Suite 1001
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       APPLICATION NUMBER: US 08/585,391
FILING DATE: 11-JAN-1996
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/576,559
FILING DATE: 21-DEC-1995
PRIOR APPLICATION DATA::
APPLICATION NUMBER: US 08/575,359
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      398 IFQEDTIPRTQIERRKTSLYFSSKYN 423
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       APPLICATION NUMBER: US/08/639,501
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 672 SSKNDTV --- CIHKDETDLLFTDQHN 694
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        COMPUTER: IBM PC compatible OPERATING SYSTEM: PC-DOS/MS-DOS
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          Sequence 2, Application US/08639501
Patent No. 5837492
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          GENERAL INFORMATION:
APPLICANT: Tavtigian, Sean V.
APPLICANT: Kamb, Alexander
APPLICANT: Simard, Jacques
APPLICANT: Couch, Fergus
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      Floppy disk
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 PRIOR APPLICATION DATA:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         COMPUTER READABLE FORM:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       Washington
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          CLASSIFICATION:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    USA
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                MEDIUM TYPE:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          ADDRESSEE:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      COUNTRY: U
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   RESULT 2
US-08-639-501-2
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               STREET:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 STATE:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            CITY:
  Matches
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                RESULT
                                                                                               Dp
                                                                                                                                                                                                                                                                                                qq
                                                                                                                                                                                                                                                                                                                                                                                               Op
                                                                                                                                                    òγ
                                                                                                                                                                                                                                                ŏ
                                                                                                                                                                                                                                                                                                                                                  δλ
                                                                                                                                                                                                                                                                                                                                                                                                                                                    ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   qq
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   ŏλ
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  В
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     QΥ
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   Ω
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      δy
```

```
PROTEINS AND SCREENING AND THERAPEUTIC METHODS BASED ON CHARACTERIZED BRCA1 AND BRCA2 PROTEINS.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        23;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         972 KSDISLNIDKIPEKNNDYMNKWAG-----LLGPISNHSFGGSFRTASNKEIKLSEHNIKK 1026
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            1027 SKMFFK-----DIEEQYPTSL---ACVEIVNTLALDNQKKLSKPQSIN-TVSAHLQSSVV 1077
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       1119 ----EESGSQFEFTQFRKPSY1LQKSTFEVPENQMTILKTTSEECRDADLHVIMNAPSIG 1174
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  1175 QVDSSKQFEGTVEIKRKFAGLLKNDCNKSASGYLTDE-----NEVGFRGFYSAHGTKL 1227
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               1228 NVSTEALQKAVKLFSDTENISEETS---AEVHPISLSSSKCHD----SVVSMF---KIEN 1277
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   1278 HNDK---TVSEKNNKCQLILQNNIEMTTGTFVEEI-----TENYKRNTENEDNKYTAA 1327
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     Gaps
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    96 QRLFGKTAGRFDVYFISPQGLKFRSKSSLANYLHKNGETSL-KPEDFDFTVLSKRGIKSR 154
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         240 ------GIPIKK-----TKKGCRKSCSGFVQSDSKRESVCNKADAESEPVAQKSQL 284
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             285 D-RTVCISDAGACGETLSVTSEENSLVKKKERSLSSGSNFCSEQKTSGIINKFCSAKDSE 343
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              344 HNEKYEDTFLESEE------IGTKVEVVERKEHLHTDILKRGSEMDNN--CSP 388
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                36 KEDVAMELERVGEDEEQMMIKRSSECNPLLQEPIASAQFGATAGTECRKSVPCGWERVVK 95
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     155 YKDCSMAALTSHL.--QNQSNNSNWNLRTRSKCKKDVFMPPSSSSELQESRGLSNFTSTHL
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     Best Local Similarity 20.9%; Pred. No. 0.0003;
Matches 96; Conservative 67; Mismatches 158; Indels 138;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            4.7%; Score 145; DB 2; Length 3418;
20.9%; Pred. No. 0.0003;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 APPLICANT: THOMPSON, MARILYN E.
TITLE OF INVENTION: CHARACTERIZED BRCA1 AND BRCA2
TITLE OF INVENTION: CHARACTERIZED BRCA1 AND THE
TITLE OF INVENTION: CHARACTERIZED BRCA1 AND BRCA2
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    :| : | | : : | | : : | | 328 SRNSHNLEFDGSDSSKNDTV---CIHKDETDLLFTDQHN 1363
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  389 TRK----DFTGEKIFQEDTIPRTQIERRKTSLYFSSKYN 423
                                                                                                                                                                                                           24884-116802-04
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           ROBINSON-BENION, CHERYL L.
                                                     JMBER: US 08/573,779
18-DEC-1995
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       ; Sequence 4, Application US/08603753D
; Patent No. 5891857
                                                                                                                                                                                                                              TELECOMMUNICATION INFORMATION:
TELEPHONE: 202-962-4810
TELEFAX: 202-962-8300
                                                                                                                                                                            28,957
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      APPLICANT: KING, MARY-CLAIRE
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  SZABO, CSILLA I.
JETTON, THOMAS L.
                                                                                                                                                                                                                                                                                                                     INFORMATION FOR SEQ ID NO: 2: SEQUENCE CHARACTERISTICS:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        APPLICANT: HOLT, JEFFREY T. APPLICANT: JENSEN, ROY A. APPLICANT: PAGE, DAVID L.
                                                                                                                ATTORNEY/AGENT INFORMATION:
                                                                                                                                                                                                                                                                                                                                                                                    3418 amino acids
20-DEC-1995
                                                                                                                                                                                                        REFERENCE/DOCKET NUMBER:
                                                                                                                                                NAME: Ihnen, Jeffrey L. REGISTRATION NUMBER: 28
                         PRIOR APPLICATION DATA:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                  MOLECULE TYPE: protein
                                                           APPLICATION NUMBER:
                                                                                                                                                                                                                                                                                                                                                                                                                   amino acid
                                                                                                                                                                                                                                                                                                                                                                                                                                                   linear
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  GENERAL INFORMATION:
                                                                                       FILING DATE:
                                                                                                                                                                                                                                                                                                                                                                                                                                         TOPOLOGY:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              US-08-603-753D-4
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     US-08-639-501-2
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  APPLICANT:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           APPLICANT:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              APPLICANT:
                                                                                                                                                                                                                                                                                                                                                                                    LENGTH:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            Query Match
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                δ
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              qq
```

```
CLASSIFICATION:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            USA
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 MEDIUM TYPE:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        FILING DATE:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        FILING DATE:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 22204
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           2
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            COUNTRY:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         STATE:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         QQ
                                                                                                                                                                                                                                                                                                                                          á
                                                                                                                                                                                                                                                                                                                                                                                                                       δλ
                                                                                                                                                                                                                                                                                                                                                                                                                                                             QQ
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     ŏ
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         OTHER INFORMATION: BRCA2 protein has a negative OTHER INFORMATION: regulatory effect on growth of human mammary cells. PUBLICATION INFORMATION:
AUTHORS: Wooster, R. et al.
TITLE: Identification of the breast cancer
TITLE: susceptability gene BRCA2
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              972 KSDISLNIDKIPEKNNDYMNKWAG-----LLGPISNHSFGGSFRTASNKEIKLSEHNIKK 1026
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           Ouery Match

4.7%: Score 145; DB 2; Length 3418;
Best Local Similarity 20.9%; Pred. No. 0.0003;
Matches 96; Conservative 67; Mismatches 158; Indels 138; Gaps
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            36 KEDVAMELERVGEDEEQMMIKRSSECNPLLQEPIASAQFGATAGTECRKSVPCGWERVVK 95
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            RELEVANT RESIDUES IN SEQ 1D NO: 4: granin box
RELEVANT RESIDUES IN SEQ 1D NO: domain at amino acids 3334-3344
                                        ADDRESSEE: ARLES A. TAYLOR, JR.
STREET: SUITE 1401, UNIVERSITY TOWER, 3100 TOWER
STREET: BOULEVARD
                                                                                                                                                                                            MEDIUM TYPE: Diskette, 3.50 inch, 800 kB storage COMPUTER: IBM PC/XT/AT compatible OPERATING SYSTEM: Windows 3.1
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               NAME/KEY: BRCA2 protein
LOCATION: 1 to 3418; Genbank locus HSU43746
IDENTIFICATION METHOD:
                                                                                                                                                                                                                                                          SOFTWARE: WORD PERFECT 6.1 and ASCII
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/603,753D
FILING DATE: 20 FEB 1996
CLASSIFICATION: 514
                                                                                                                                                                                                                                                                                                                                                         PRIOR APPLICATION DATA:
APPLICATION NUMBER: U.S. 08/373,799
FILING DATE: 17 JAN 1995
ATTORNEY/AGENT INFORMATION:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        1242/2
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           DEVELOPMENTAL STAGE: adult
TISSUE TYPE: female breast
CELL TYPE: normal breast tissue
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         ORGANISM: Homo sapiens sapiens
INDIVIDUAL ISOLATE:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    REFERENCE/DOCKET NUMBER: 124
TELECOMMUNICATION INFORMATION:
                                                                                                                                                                                                                                                                                                                                                                                                                                        NAME: ARLES A. TAYLOR, JR. REGISTRATION NUMBER: 39,395
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          TELEPHONE: (919) 493-8000
TELEFAX: (919) 419-0383
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         SEQUENCE CHARACTERISTICS: LENGTH: 3418
                                                                                                                      NORTH CAROLINA
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  :SS: single
unknown
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  INFORMATION FOR SEQ ID NO:
                     CORRESPONDENCE ADDRESS:
                                                                                                                                                                             COMPUTER READABLE FORM:
MEDIUM TYPE: Diskett
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    MOLECULE TYPE: protein HYPOTHETICAL: no
NUMBER OF SEQUENCES:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 TYPE: amino acid
STRANDEDNESS: sin
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                JOURNAL: Nature
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             HMEC
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      789-792
                                                                                                                                              USA
                                                                                                   DURHAM
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 ANTI-SENSE: no
ORIGINAL SOURCE:
                                                                                                                                                          27707
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      TELEPHONE:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        CELL LINE:
ORGANELLE:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             1995
                                                                                                                                       COUNTRY:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         ORGANISM:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    US-08-603-753D-4
                                                                                                                      STATE:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        PAGES:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                q
```

```
1027 SKMFFK----DIEEQYPTSL---ACVEIVNTLALDNOKKLSKPQSIN-TVSAHLQSSVV 1077
                                                                                                                                   1078 VSDCKNSHITPQMLFSKQDFNSNHNL-----TPSQKAEITE-----LSTIL 1118
                                                                                                                                                                                                                               :| | : || | 119 ----EESGSQFEFTQFRKPSYILORSTFEVPENOMTILKTTSEECRDADLHVIMNAPSIG 1174
                                                                                                                                                                                                                                                                                                                                1175 QVDSSKQFEGTVEIKRRFAGLLKNDCNKSASGYLTDE-----NEVGFRGFYSAHGTKL 1227
                                                                                                                                                                                                                                                                                                                                                                                                                               1228 NVSTEALQKAVKLESDIENISEETS---AEVHPISLSSSKCHD----SVVSMF---KIEN 1277
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                155 YKDCSMAALTSHL--QNQSNNSNWNLRTRSKCKKDVFMPPSSSSELQESRGLSNFTSTHL 212
                                                                                                                                                                                               96 QRLFGKTAGRFDVYFISPQGLKFRSKSSLANYLHKNGETSL-KPEDFDFTVLSKRGIKSR 154
                                                                                                                                                                                                                                                                                              240 -----GIPIKK----TKKGCRKSCSGFVQSDSKRESVCNKADAESEPVAQKSQL 284
                                                                                                                                                                                                                                                                                                                                                                                             285 D-RTVCISDAGACGETLSVTSEENSLVKKKERSLSSGSNFCSEQKTSGIINKFCSAKDSE 343
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              344 HNEKYEDTFLESEE-----GTRVEVERVERKEHLHTDILKRGSEMDNN--CSP 388
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     TITLE OF INVENTION: Chromosome 13-Linked Breast Cancer FITLE OF INVENTION: Susceptibility Gene NUMBER OF SEQUENCES: 124
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   ADDRESSEE: Venable, Baetjer, Howard & Civiletti
STREET: 1201 New York Avenue N.W., Suite 1001
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  1328 SRNSHNLEFDGSDSSKNDTV---CIHKDETDLLFTDQHN 1363
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                SOFTWARE: Patentin Release #1.0, Version #1.30 CURRENT APPLICATION DATA:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  389 TRK----DFTGEKIFQEDTIPRTQIERRKTSLYFSSKYN 423
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       APPLICATION NUMBER: US/09/044,946
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/576,559
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            US 08/575,359
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   US 08/573,779
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   PC-DOS/MS-DOS
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                US-09-044-946-2; Sequence 2, Application US/09044946; Patent No. 6033857
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              PRIOR APPLICATION DATA:
APPLICATION NUMBER: 08/639,501
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          GENERAL INFORMATION:
APPLICANT: Tavtigian, Sean V.
APPLICANT: Kamb, Alexander
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      COMPUTER: IBM PC COmpatible OPERATING SYSTEM: PC-DOS/MS-
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         APPLICANT: Simard, Jacques
APPLICANT: Couch, Fergus
APPLICANT: Rommens, Johanna
APPLICANT: Weber, Barbara
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     FILING DATE: 21-DEC-1995
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08//
FILING DATE: 20-DEC-1995
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08//
FILING DATE: 18-DEC-1995
ATTORNEY/AGENT INFORMATION:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 Floppy disk
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          lhnen, Jeffrey L.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         CORRESPONDENCE ADDRESS:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         COMPUTER READABLE FORM:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        Washington
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            NAME:
```

```
RESULT
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 pp
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  qq
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            TITLE OF INVENTION: Materials and methods relating to the TITLE OF INVENTION: identification and sequencing of the BRCA2 cancer TITLE OF INVENTION: susceptibility gene and uses thereof.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     1228 NVSTEALQKAVKLESDIENISEETS---AEVHPISLSSSKCHD----SVVSMF---KIEN 1277
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           1278 HNDK---TVSEKNNKCOLILQNNIEMTTGTFVEEL-----TENYKRNTENEDNKYTAA 1327
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  972 KSDISLNIDKIPEKNNDYMNKWAG-----LLGPISNHSFGGSFRTASNKEIKLSEHNIKK 1026
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              1027 SKMFFK-----DIEEQYPTSL---ACVEIVNTLALDNQKKLSKPQSIN-TVSAHLQSSVV 1077
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               1078 VSDCKNSHITPQMLFSKQDFNSNHNL------TPSOKAEITE-----LSTIL 1118
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                1119 ----EESGSQFEFTQFRKPSYILQKSTFEVPENOMTILKTTSEECRDADLHVIMNAPSIG 1174
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         1175 QVDSSKQFEGTVEIKRKFAGLLKNDCNKSASGYLTDE-----NEVGFRGFYSAHGTKL 1227
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             285 D-RIVCISDAGACGETLSVTSEENSLVKKKERSLSSGSNFCSEQKTSGIINKFCSAKDSE 343
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              344 HNEKYEDTFLESEE------GTKVEVVERKEHLHTDILKRGSEMDNN--CSP 388
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                155 YKDCSMAALTSHL--QNQSNNSNWNLRTRSKCKKDVFMPPSSSSELQESRGLSNFTSTHL 212
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               240 ------GIPIKK----TKKGCRKSCSGFVQSDSKRESVCNKADAESEPVAQKSQL 284
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 96 QRLFGKTAGRFDVYFISPQGLKFRSKSSLANYLHKNGETSL-KPEDFDFTVLSKRGIKSR 154
                                                                                                                                                                                                                                                                                                                               Query Match
4.7%; Score 145; DB 3; Length 3418;
Best Local Similarity 20.9%; Pred. No. 0.0003;
Matches 96; Conservative 67; Mismatches 158; Indels 138; Gaps
                                                                                                                                                                                                                                                                                                                                                                                                                                   36 KEDVAMELERVGEDBEÇMMIKRSSECNPLLQEPIASAQFGATAGTECRKSVPCGWERVVK 95
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       E: Bell Seltzer Park & Gibson
310 UCB Plaza, 3605 Glenwood Avenue, PO Drawer 31107
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      213 LLKEDEGVDDVNFRKVRK-------PKGKVTILK-----
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               1328 SRNSHNLEFDGSDSSKNDTV---CIHKDETDLLFTDQHN 1363
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  389 TRK----DFTGEKIFQEDTIPRTQIERRKTSLYFSSKYN 423
                    24884-116802-04
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            Sequence 44, Application US/08755587 Patent No. 6045997
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      APPLICANT: Futreal, Phillip A
APPLICANT: Wooster, Richard F
APPLICANT: Ashworth, Alan
APPLICANT: Stratton, Michael R
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          COMPUTER: IBM PC compatible
                    REFERENCE/DOCKET NUMBER: 248
TELECOMMUNICATION INFORMATION:
TELEPHONE: 202-962-4810
28,957
                                                                                                                   INFORMATION FOR SEQ ID NO: 2:
SEQUENCE CHARACTERISTICS:
LENGTH: 3418 amino acids
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  Floppy disk
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        COMPUTER READABLE FORM:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          CORRESPONDENCE ADDRESS:
                                                                                              202-962-8300
                                                                                                                                                                                                                                                 MOLECULE TYPE: protein
  REGISTRATION NUMBER:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                NUMBER OF SEQUENCES:
                                                                                                                                                                                                    amino acid
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   GENERAL INFORMATION:
                                                                                                                                                                                                                         linear
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    Raleigh
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  MEDIUM TYPE:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     ADDRESSEE:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          US-08-755-587-44
                                                                                                                                                                                                                         TOPOLOGY:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   COUNTRY:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            STREET:
                                                                                                      TELEFAX:
                                                                                                                                                                                                                                                                        US-09-044-946-2
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        pp
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            QΥ
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           qq
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             δy
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            q
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               qq
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 δλ
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               δλ
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               δλ
```

```
1278 HNDK---TVSEKNNKCQLILQNNIEMTTGTFVBEI-----TENYKNTENBDNKYTAA 1327
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              1027 SKMFFK-----DIEEQYPTSL---ACVEIVNTLALDNOKKLSKPQSIN-TVSAHLQSSVV 1077
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   1078 VSDCKNSHITPQMLFSKQDFNSNHNL-----TPSQKAEITE-----LSTIL 1118
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              1119 ----EESGSQFEFTQFRKPSY1LQKSTFEVPENQMTILKTTSEECRDADLHVIMNAPSIG 1174
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          : | : : | : : | : : | : : | : : | : : | : : | : : | : : | : : | : : | : : | : : | : : | : : | : : | : : | : : | : : | : : | : : | : : | : : | : : | : : | : : | : : | : : | : : | : : | : : | : : | : : | : : | : : | : : | : : | : : | : : | : : | : : | : : | : : | : : | : : | : : | : : | : : | : : | : : | : : | : : | : : | : : | : : | : : | : : | : : | : : | : : | : : | : : | : : | : : | : : | : : | : : | : : | : : | : : | : : | : : | : : | : : | : : | : : | : : | : : | : : | : : | : : | : : | : : | : : | : : | : : | : : | : : | : : | : : | : : | : : | : : | : : | : : | : : | : : | : : | : : | : : | : : | : : | : : | : : | : : | : : | : : | : : | : : | : : | : : | : : | : : | : : | : : | : : | : : | : : | : : | : : | : : | : : | : : | : : | : : | : : | : : | : : | : : | : : | : : | : : | : : | : : | : : | : : | : : | : : | : : | : : | : : | : : | : : | : : | : : | : : | : : | : : | : : | : : | : : | : : | : : | : : | : : | : : | : : | : : | : : | : : | : : | : : | : : | : : | : : | : : | : : | : : | : : | : : | : : | : : | : : | : : | : : | : : | : : | : : | : : | : : | : : | : : | : : | : : | : : | : : | : : | : : | : : | : : | : : | : : | : : | : : | : : | : : | : : | : : | : : | : : | : : | : : | : : | : : | : : | : : | : : | : : | : : | : : | : : | : : | : : | : : | : : | : : | : : | : : | : : | : : | : : | : : | : : | : : | : : | : : | : : | : : | : : | : : | : : | : : | : : | : : | : : | : : | : : | : : | : : | : : | : : | : : | : : | : : | : : | : : | : : | : : | : : | : : | : : | : : | : : | : : | : : | : : | : : | : : | : : | : : | : : | : : | : : | : : | : : | : : | : : | : : | : : | : : | : : | : : | : : | : : | : : | : : | : : | : : | : : | : : | : : | : : | : : | : : | : : | : : | : : | : : | : : | : : | : : | : : | : : | : : | : : | : : | : : | : : | : : | : : | : : | : : | : : | : : | : : | : : | : : | : : | : : | : : | : : | : : | : : | : : | : : | : : | : : | : : | : : | : : | : : | : : | : : | : : | : : | : : | : : | : : | : : | : : | : : | : : | : : | : : | : : | : : | : : | : : | : : | : : | : 
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             1228 NVSTEALQKAVKLFSDIENISETS---AEVHPISLSSSKCHD----SVVSMF---KIEN 1277
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         972 KSDISLNIDKIPEKNNDYMNKWAG-----LLGPISNHSFGGSFRTASNKEIKLSEHNIKK 1026
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   285 D-RTVCISDAGACGETLSVTSEENSLVKKKERSLSSGSNFCSEQKTSGIINKFCSAKDSE 343
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             344 HNEKYEDTFLESEE------GTKVEVVERKEHLHTDILKRGSEMDNN--CSP 388
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         155 YKDCSMAALTSHL--QNQSNNSNWNIRTRSKCKKDVFMPPSSSSELQESRGLSNFTSTHL 212
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             240 ------GIPIKK----TKKGCRKSCSGFVQSDSKRESVCNKADAESEPVAQKSQL 284
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   96 QRLFGKTAGRFDVYFISPQGLKFRSKSSLANYLHKNGETSL·KPEDFDFTVLSKRGIKSR 154
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    Oucry Match 4.7%; Score 145; DB 3; Length 3418;
Best Local Similarity 20.9%; Pred. No. 0.0003;
Matches 96; Conservative 67; Mismatches 158; Indels 138; Gaps
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              36 KEDVAMELERVGEDEEQMMIKRSSECNPLLQEPIASAQFGATAGTECRKSVPCGWERVVK 95
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        213 ILKEDEGVDDVNFRKVRK------PKGKVTILK-----
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentin Release #1.0, Version #1.25 (EPO)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              389 TRK----DFTGEKIFQEDTIPRTQIERRKTSLYFSSKYN 423
                                                                                                                                  US/08/755,587
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 UMBER: GB 9617961.9
28-AUG-1996
                                                                                                                                                                                                                             PRIOR APPLICATION DATA:
APPLICATION NUMBER: GB 9523959.6
FILING DATE: 23-NOV-1995
                                                                                                                                                                                                                                                                                                                                                                                                             JMBER: GB 9525555.0
14-DEC-1995
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    ; Sequence 2, Application US/09044908
; Patent No. 6124104
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      Tavtigian, Sean V.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    /AGENT LINE.
Kenneth D Sibley 31,665
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     44:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         Kamb, Alexander
Simard, Jacques
Couch, Fergus
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     ATTORNEY/AGENT INFORMATION:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               3418 amino acids
                                                                                                                                                                           25-NOV-1996
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             REFERENCE/DOCKET NUMBER:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  SEQUENCE CHARACTERISTICS:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          INFORMATION FOR SEQ ID NO:
                                                                                CURRENT APPLICATION DATA: APPLICATION NUMBER: US
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            MOLECULE TYPE: protein
                                                                                                                                                                                                                                                                                                                                                                     PRIOR APPLICATION DATA:
APPLICATION NUMBER:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             PRIOR APPLICATION DATA:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     REGISTRATION NUMBER:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          APPLICATION NUMBER:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       amino acid
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   GENERAL INFORMATION:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    linear
                                                                                                                                                                                                                                                                                                                                                                                                                                                                          FILING DATE:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       FILING DATE:
                                                                                                                                                                                          FILING DATE:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     APPLICANT:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   APPLICANT:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                APPLICANT:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             APPLICANT:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           US-09-044-908-2
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               TOPOLOGY:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          US-08-755-587-44
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              LENGTH:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       NAME:
```

```
APPLICATION NUMBER: U.S. 08/373,799
FILING DATE: 17 JAN 1995
ATTORNEY/AGENT INFORMATION:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                FILING DOLL.
CLASSIFICATION:
PRIOR APPLICATION DATA:
                                                                                                                                                                                                                                                                                                           ; Sequence 4, Application US/09099753; Patent No. 6149903
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         Windows 3.1
                                                                                                                                                                                                                                                                                                                                                                                     JEGS, OSTING A.
PAGE, DAVID L.
KING, MARY-CLAIRE
SZABO, CSILLA I.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 ARLES A. TAYLOR, JR.
RATION NUMBER: 39,395
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        REFERENCE/DOCKET NUMBER: 12/
TELECOMMUNICATION INFORMATION:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                  JETTON, THOMAS L.
                                                                                                                                                                                                                                                                                                                                                                       HOLT, JEFFREY T.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               TELEPHONE: (919) 493-8000
TELEFAX: (919) 419-0383
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        FILING DATE: 20 FEB 1996
APPLICATION NUMBER: U.S.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       NORTH CAROLINA
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          SEQUENCE CHARACTERISTICS:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  INFORMATION FOR SEQ ID NO:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       CORRESPONDENCE ADDRESS:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            COMPUTER READABLE FORM:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      SS: single
unknown
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  MOLECULE TYPE: protein
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        REGISTRATION NUMBER:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     NUMBER OF SEQUENCES:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           TITLE OF INVENTION:
TITLE OF INVENTION:
TITLE OF INVENTION:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     OPERATING SYSTEM:
                                                                                                                                                                                                                                                                                                                                                  GENERAL INFORMATION:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   amino acid
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           HYPOTHETICAL: no
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           USA
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    DURHAM
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           ORIGINAL SOURCE:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  MEDIUM TYPE:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               3418
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 STRANDEDNESS:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             TELEPHONE:
                                                                                                                                                                                                                                                                                                                                                                                                         APPLICANT:
APPLICANT:
                                                                                                                                                                                                                                                                                     US-09-099-753-4
                                                                                                                                                                                                                                                                                                                                                                                         APPLICANT:
                                                                                                                                                                                                                                                                                                                                                                                                                                                APPLICANT:
                                                                                                                                                                                                                                                                                                                                                                  APPLICANT:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                    APPLICANT:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         APPLICANT:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         COUNTRY:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   COMPUTER:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               ANTI-SENSE:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             LENGTH:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       STATE:
                                                                                                                                                                                                                                                                       RESULT
                                                                                                                                  Op
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            972 KSDISLNIDKIPEKNNDYMNKWAG-----LLGPISNHSFGGSFRTASNKEIKLSEHNIKK 1026
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        1027 SKMFFK-----DIEEQYPTSL---ACVEIVNTLALDNQKKLSKPQSIN-TVSAHLQSSVV 1077
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      :119 ----EESGSQFEFTQFRKPSYILQKSTFEVPENQMTILKTTSEECRDADLHVIMNAPSIG 1174
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         1175 QVDSSKQFEGTVEIKRKFAGLLKNDCNKSASGYLTDE-----NEVGFRGFYSAHGTKL 1227
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             Query Match
4.7%: Score 145; DB 3; Length 3418;
Best Local Similarity 20.9%; Pred. No. 0.0003;
Matches 96: Conservative 67; Mismatches 158; Indels 138; Gaps
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                155 YKDCSMAALTSHL--QNQSNNSNWNLRTRSKCKKDVFMPPSSSSELQESRGLSNFTSTHL 212
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    96 QRLFGKTAGRFDVYFISPQGLKFRSKSSLANYLHKNGETSL-KPEDFDFTVLSKRGIKSR 154
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       240 ------GIPIKK----TKKGCRKSCSGFVQSDSKRESVCNKADAESEFVAQKSQL 284
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           36 KEDVAMELERVGEDEEQMMIKRSSECNPLLQEPIASAQFGATAGTECRKSVPCGWERVVK 95
                                    TITLE OF INVENTION: Chromosome 13-Linked Breast Cancer TITLE OF INVENTION: Susceptibility Gene NUMBER OF SEQUENCES: 124
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           213 LLKEDEGVDDVNFRKVRK------PKGKVTILK------
                                                                                       CORRESPONDENCE ADDRESS:
ADDRESSEE: Venable, Baetjer, Howard & Civiletti
STREET: 1201 New York Avenue N.W., Suite 1001
                                                                                                                                                                                                                        COMPUTER READBLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentin Release #1.0, Version #1.30
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       24884-116802-04
                                                                                                                                                                                                                                                                                                                                                                                                                                                        PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/576,559
FILING DATE: 21-DEC-1995
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/575,359
FILING DATE: 20-DEC-1995
PRIOR APPLICATION NUMBER: US 08/573,779
FILING DATE: 18-DEC-1995
ATTORNEY AGENT INFORMATION:
                                                                                                                                                                                                                                                                                                                                                APPLICATION NUMBER: US/09/044,908
                                                                                                                                                                                                                                                                                                                                                                                                     PRIOR APPLICATION DATA: APPLICATION NUMBER: 08/639,501
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        NAME: Ihnen, Jeffrey L. REGISTRATION NUMBER: 28,957
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   REFERENCE/DOCKET NUMBER: 246
TELECOMMUNICATION INFORMATION:
TELEPHONE: 202-962-4810
TELEFAX: 202-962-8300
APPLICANT: Rommens, Johanna
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      TELEFAX: 202-962-8300
INFORMATION FOR SEQ ID NO: 2:
                      Weber, Barbara
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   3418 amino acids
                                                                             124
                                                                                                                                                                                                                                                                                                                            CURRENT APPLICATION DATA:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             SEQUENCE CHARACTERISTICS:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  ; MOLECULE TYPE: protein US-09-044-908-2
                                                                                                                                                    Washington
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     amino acid
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            linear
                                                                                                                                                                                                                                                                                                                                                                                       CLASSIFICATION:
                                                                                                                                                                                             USA
                                                                                                                                                                                                                                                                                                                                                                                                                                              FILING DATE:
                                                                                                                                                                                                              22204
                                                                                                                                                                                                                                                                                                                                                                  FILING DATE
                                                                                                                                                                                             COUNTRY:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      TOPOLOGY:
                                                                                                                                                                       STATE
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             ò
```

```
THOMPSON, MARILYN E.
VEWIION: CHARACTERIZED BRCA1 AND BRCA2
VENTION: PROTEINS AND SCREENING AND THERAPEUTIC METHODS BASED ON
VENTION: CHARACTERIZED BRCA1 AND BRCA2 PROTEINS.
                                                                                                                                                                                      1228 NVSTEALQKAVKLFSDIENISEETS---AEVHPISLSSSKCHD----SVVSMF---KIEN 1277
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  1278 HNDK---TVSEKNNKCQLILQNNIEMTTGTFVEEI-----TENYKRNTENEDNKYTAA 1327
285 D-RTVCISDAGACGETLSVTSEENSLVKKKERSLSSGSNFCSEQKTSGIINKFCSAKDSE 343
                                                                                                                                                                                                                                                                                                                                                                                      344 HNEKYEDTFLESEE-----CGTKVEVVERKEHLHTDILKRGSEMDNN--CSP 388
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       ADDRESSEE: ARLES A. TAYLOR, JR.
STREET: SUITE 1401, UNIVERSITY TOWER, 3100 TOWER
STREET: BOULEVARD
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    E: Diskette, 3.50 inch, 800 kB storage IBM PC/XT/AT compatible
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            :| :| | ::| | ::| | ::| | ::| | ::| | ::| | ::| | ::| | ::| | ::| | ::| | ::| | ::| | ::| | ::| | ::| | ::| | ::| | ::| | ::| | ::| | ::| | ::| | ::| | ::| | ::| | ::| | ::| | ::| | ::| | ::| | ::| | ::| | ::| | ::| | ::| | ::| | ::| | ::| | ::| | ::| | ::| | ::| | ::| | ::| | ::| | ::| | ::| | ::| | ::| | ::| | ::| | ::| | ::| | ::| | ::| | ::| | ::| | ::| | ::| | ::| | ::| | ::| | ::| | ::| | ::| | ::| | ::| | ::| | ::| | ::| | ::| | ::| | ::| | ::| | ::| | ::| | ::| | ::| | ::| | ::| | ::| | ::| | ::| | ::| | ::| | ::| | ::| | ::| | ::| | ::| | ::| | ::| | ::| | ::| | ::| | ::| | ::| | ::| | ::| | ::| | ::| | ::| | ::| | ::| | ::| | ::| | ::| | ::| | ::| | ::| | ::| | ::| | ::| | ::| | ::| | ::| | ::| | ::| | ::| | ::| | ::| | ::| | ::| | ::| | ::| | ::| | ::| | ::| | ::| | ::| | ::| | ::| | ::| | ::| | ::| | ::| | ::| | ::| | ::| | ::| | ::| | ::| | ::| | ::| | ::| | ::| | ::| | ::| | ::| | ::| | ::| | ::| | ::| | ::| | ::| | ::| | ::| | ::| | ::| | ::| | ::| | ::| | ::| | ::| | ::| | ::| | ::| | ::| | ::| | ::| | ::| | ::| | ::| | ::| | ::| | ::| | ::| | ::| | ::| | ::| | ::| | ::| | ::| | ::| | ::| | ::| | ::| | ::| | ::| | ::| | ::| | ::| | ::| | ::| | ::| | ::| | ::| | ::| | ::| | ::| | ::| | ::| | ::| | ::| | ::| | ::| | ::| | ::| | ::| | ::| | ::| | ::| | ::| | ::| | ::| | ::| | ::| | ::| | ::| | ::| | ::| | ::| | ::| | ::| | ::| | ::| | ::| | ::| | ::| | ::| | ::| | ::| | ::| | ::| | ::| | ::| | ::| | ::| | ::| | ::| | ::| | ::| | ::| | ::| | ::| | ::| | ::| | ::| | ::| | ::| | ::| | ::| | ::| | ::| | ::| | ::| | ::| | ::| | ::| | ::| | ::| | ::| | ::| | ::| | ::| | ::| | ::| | ::| | ::| | ::| | ::| | ::| | ::| | ::| | ::| | ::| | ::| | ::| | ::| | ::| | ::| | ::| | ::| | ::| | ::| | ::| | ::| | ::| | ::| | ::| | ::| | ::| | ::| | ::| | ::| | ::| | ::| | ::| | ::| | ::| | ::| | ::| | ::| | ::| | ::| | ::| | ::| | ::| | ::| | ::| | ::| | ::| | ::| | ::| | ::| | ::| | ::| | ::| | ::| | ::| | ::| | ::| | ::| | ::| | ::| | ::| | ::| | ::| | ::| | ::| | ::| | ::| | ::| | ::| | ::| | ::| | ::| | ::| | ::| 
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            389 TRK----DFTGEKIFQEDTIPRTQIERRKTSLYFSSKYN 423
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       SOFTWARE: WORD PERFECT 6.1 and ASCII CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/09/099,753
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      APPLICANT: ROBINSON-BENION, CHERYL L.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     ORGANISM: Homo sapiens sapiens
INDIVIDUAL ISOLATE:
DEVELOPMENTAL STAGE: adult
```

```
OTHER INFORMATION: BRCA2 protein has a negative OTHER INFORMATION: regulatory effect on growth of human mammary cells.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          1228 NVSTEALQKAVKLFSDIENISEETS---AEVHPISLSSSKCHD----SVVSMF---KIEN 1277
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        1278 HNDK---TVSEKNNKCQLILQNNIEMTTGTFVEEI-----TENYKRNTENEDNKYTAA 1327
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         1078 VSDCKNSHITPQMLFSKQDFNSNHNL-----TPSQKAEITE-----LSTIL 1118
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         1175 QVDSSKQFEGTVEIKRKFAGLLKNDCNKSASGYLTDE-----NEVGFRGFYSAHGTKL 1227
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          1119 ----EESGSQFEFTQFRKPSYILQKSTFEVPENQMTILKTTSEECRDADLHVIMNAPSIG 1174
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           1027 SKMFFK-----DIEEQYPTSL---ACVEIVNTLALDNQKKLSKFQSIN-TVSAHLQSSVV 1077
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          285 D-RTVCISDAGACGETLSVTSEENSLVKKKERSLSSGSNFCSEQKTSGIINKFCSAKDSE 343
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            344 HNEKYEDTFLESEE------IGTKVEVVERVERKEHLHTDILKRGSEMDNN--CSP 388
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            155 YKDCSMAALTSHL--QNQSNNSNWNLRTRSKCKKDVFMPPSSSSELQESRGLSNFTSTHL 212
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           213 LLKEDEGVDDVNFRKVRK------239
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         240 ------GIPIKK-----TKKGCRKSCSGFVQSDSKRESVCNKADAESEPVAQKSQL 284
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           96 QRLFGKTAGRFDVYFISPQGLKFRSKSSLANYLHKNGETSL-KPEDFDFTVLSKRGIKSR 154
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  Gaps
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            36 KEDVAMELERVGEDEEQMMIKRSSECNPLLQEPIASAQFGATAGTECRKSVPCGWERVVK 95
                                                                                                                                                                                                                                                                                                                                                                                                                    RELEVANT RESIDUES IN SEQ ID NO: 4: granin box RELEVANT RESIDUES IN SEQ ID NO: domain at amino acids 3334-3344 US-09-099-753-4
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           Ouery Match
Best Local Similarity 20.9%; Pred. No. 0.0003;
Matches 96; Conservative 67; Mismatches 158; Indels 138;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               :| : :| | : ::| | : ::| | 328 SRNSHNLEFDGSDSSKNDTV---CIHKDETDLLFTDQHN 1363
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                389 TRK----DFTGEKIFQEDTIPRTQIERRKTSLYFSSKYN 423
                                                                                                                 NAME/KEY: BRCA2 protein
LOCATION: 1 to 3418; Genbank locus HSU43746
IDENTIFICATION METHOD:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      APPLICANT: STEINER, MITCHELL S.
APPLICANT: ROBINSON-BENION, CHERYL L.
APPLICANT: THOMROON, MARILYN E.
TITLE OF INVENTION: THERAPECTIC METHODS FOR
TITLE OF INVENTION: PROSTATE CANCER
                                                                                                                                                                                                                  OTHER INFORMATION: regulatory effect on gro
PUBLICATION INFORMATION:
AUTHORS: Wooster, R. et al.
AUTHORS: Gooster, R. et al.
Identification of the breast cancer
TITLE: susceptability gene BRCA2
TISSUE TYPE: female breast
CELL TYPE: normal breast tissue
CELL LINE: HMEC
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           Sequence 4, Application US/08986106 Patent No. 6177410
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    KING, MARY-CLAIRE
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    HOLT, JEFFREY T.
JENSEN, ROY A.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               GENERAL INFORMATION:
APPLICANT: HOLT, J
APPLICANT: JENSEN,
                                                                                                                                                                                                                                                                                                                                                JOURNAL: Nature
                                                                                                                                                                                                                                                                                                                                                                   VOLUME: 379
PAGES: 789-792
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      APPLICANT:
APPLICANT:
APPLICANT:
                                                                            ORGANELLE:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      RESULT 8
US-08-986-106-4
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 δλ
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           QQ
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               δλ
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              δλ
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                δλ
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            g
```

```
1078 VSDCKNSHITPQMLFSKQDFNSNHNL------TPSOKAEITE-----LSTIL 1118
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  1175 QVDSSKQFEGTVEIKRRFAGLLKNDCNKSASGYLTDE------NEVGFRGFYSAHGTML 1227
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     972 KSDISLNIDKIPEKNNDYMNKWAG-----LEGPISNHSFGGSFRTASNKEIKLSEHNIKK 1026
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               1027 SKMFFK-----DIEEQYPTSL---ACVEIVNTLALDNQKKLSKPQSIN-TVSAHLQSSVV 1077
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   1119 ----EESGSQFEFTQFRKPSYILOKSTFEVPENOMTILKTTSEECRDADLHVIMNAPSIG 1174
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       155 YKDCSMAALTSHL--QNQSNNSNWNLRTRSKCKKDVFMPPSSSSELQESRGLSNFTSTHL 212
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           240 ------GIPIKK----TKKGCRKSCSGFVQSDSKRESVCNKADAESEPVAQKSQL 284
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      96 QRLFGKTAGRFDVYFISPQGLKFRSKSSLANYLHKNGETSL-KPEDFDFTVLSKRGIKSR 154
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     36 KEDVAMELERVGEDEEQMMIKRSSECNPLLQEPIASAQFGATAGTECRKSVPCGWERVVK 95
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  4 7%; Score 145; DB 4; Length 3418;
ilarity 20.9%; Pred. No. 0.0003;
Conservative 67; Mismatches 158; Indels 138;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           RELEVANT RESIDUES IN SEQ ID NO: 4: granin box domain at RELEVANT RESIDUES IN SEQ ID NO: amino acids 3334-3344
                                                                                                                                                                                                   E: Diskette, 3.50 inch, 1.44MB storage IBM PC/XT/AT compatible
                                 E: ARLES A. TAYLOR, JR. SUITE 1401, UNIVERSITY TOWER, 3100 TOWER
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   Identification of the breast cancer susceptability gene BRCA2
                                                                                                                                                                                                                                                                SOFTWARE: WORD PERFECT 6.1 and ASCII
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/986,106
                                                                                                                                                                                                                                                                                                                                                                                                                                                           NAME: ARLES A. TAYLOR, JR. REGISTRATION NUMBER: 39,395
REPERENCE/DOCKET NUMBER: 1242/3
TELECOMMUNICATION INPORMATION:
                                                                                                                                                                                                                                                                                                                                                                            PRIOR APPLICATION DATA:
APPLICATION NUMBER: 08/603,753
FILING DATE: 20 FEB 1996
ATTORNEY/AGENT INFORMATION:
                                                                                                                                                                                                                                                     OPERATING SYSTEM: Windows 3.1
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        Wooster, R. et al.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              TELEPHONE: (919) 493-8000
TELEFAX: (919) 419-0383
INFORMATION FOR SEQ ID NO: 4:
SEQUENCE CHARACTERISTICS:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           NAME/KEY: LRCA2 protein PUBLICATION INFORMATION:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 STRANDEDNESS: single
TOPOLOGY: unknown
                                                                                                                           NORTH CAROLINA
                                                                                                                                                                                        COMPUTER READABLE FORM:
NUMBER OF SEQUENCES: 20
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  Local Similarity
hes 96; Conserva
                                                                                   BOULEVARD
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               amino acid
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       Nature
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           789-792
                                                                                                                                                                                                                                                                                                                                                            CLASSIFICATION:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        3.418
                                                                                                                                                USA
                                                                                                         DURHAM
                                                                                                                                                                                                                 MEDIUM TYPE:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  1995
                                                                                                                                                                                                                                                                                                                                             FILING DATE:
                                                                                                                                                               27707
                                             ADDRESSEE:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       JOURNAL:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        AUTHORS:
                                                                                                                                                                                                                                     COMPUTER:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              VOLUME:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   US-08-986-106-4
                                                                                                                                                  COUNTRY:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                Query Match
                                                                 STREET:
                                                                                       STREET:
                                                                                                                       STATE:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             Matches
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   q
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           δ
```

Oy Dp	285 D-RTVCISDAGACGETLSVTSEENSLVKKRERSLSSGSNFCSEQKTSGIINKFCSAKDSE 343 :	
QQ .		
Oy Dp	389 TRKDFTGEKIFQEDTIPRTQIERRKTSLYFSSKYN 423 	
RESULT US-08- ; Sequ ; Pate ; GEN	SSULT 9 5.08-755-587-16 Sequence 16, Application US/08755587 Sequence 10, 6045997 GENERAL INFORMATION: APPLICANT: Futreal, Phillip A APPLICANT: Wooster. Richard E	
	Chard Nan Michael Merial Mertifi Scepti	
	CORRESPONDENCE ADDRESS: ADDRESSE: Bell Seltzer Park & Gibson STREET: 310 UCB Plaza, 3605 Glenwood Avenue, PO Drawer 31107 CITY: Raleigh STATE: NC	
	Release #	
	US/U8/755,5 V-1996 A:	
	GB 9323555 C-1995 A:	
	FILING DATE: 28-A0(-1996 ATTORNEY STORMER: 28-A0(-1996 ATTORNEY STORMEY STORMEY STORMEY STORMEY STORMEY STORMEY STORMEY STORMEY STORMER STORME	
۶ 	SIDIEY (BER: 3	
	SEQUENCE CHARACTERISTICS: LENGINGE CHARACTERISTICS: TWORL	
; ; us-08	Tre: autho actd TOPOLOGY: linear MOLECULE TYPE: protein -08-755-587-16	
Quer Best Matcl	<pre>Query Match Best Local Similarity 21.1%; Pred. No. 0.0002; Matches 97; Conservative 66; Mismatches 158; Indels 138; Gaps 24;</pre>	
Oy Db	36 KEDVAMELERVGEDEEQMMIKRSSECNPLLQEPIASAQFGATAGTECRKSVPCGWERVVK 95	
ογ	15	
QQ		

```
TITLE OF INVENTION: TAT-SF: Cofactor For Stimulation of Transcriptional TITLE OF INVENTION: Elongation By HIV-1 TAT TITLE OF INVENTION: Elongation By HIV-1 TAT CURRENT ELEMBERINE: M056/70042 CURRENT APPLICATION NUMBER: US/09/214,564A PRIOR APPLICATION NUMBER: US 60/021,218 PRIOR PRIOR APPLICATION NUMBER: US 60/021,218 PRIOR FILING DATE: 1996-07-03 PRIOR FILING DATE: 1996-07-13 PRIOR FILING DATE: 1996-07-13 PRIOR FILING DATE: 1996-12-13 PRIOR FILING DATE: 1997-07-03 PRIOR FILING DATE: US 60/033,152 PRIO
155 YKDCSMAALTSHL--QNQSNNSNWNLRTRSKCKKDVFMPPSSSSELQESRGLSNFTSTHL 212
                                                       894 QVDSSKQFEGTVEIKRKFAGLLKNDCNKSASGYLTDE-----NEVGFRGFYSAHGTKL 946
                                                                                                                                                                                    213 LLKEDEGVDDVNFRKVRK-------PKGKVTILK--------239
                                                                                                                                                                                                                                                                         838 ---EDSG-SQFEFTQFRKPSYILQKSTFEVPENQMTILKTTSEECRDADLHVIMNAPSIG 893
                                                                                                                                                                                                                                                                                                                                                                             240 -----GIPIKK----TKKGCRKSCSGFVQSDSKRESVCNKADAESEPVAQKSQL 284
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  285 D-RTVC1SDAGACGETLSVTSEENSLVKKKERSLSSGSNFCSBQKTSGIINKFCSAKDSE 343
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              947 NVSTEALQKAVKLESDIENISEETS---AEVHPISLSSSKCHD----SVVSMF---KIEN 996
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           344 HNEKYEDTFLESEE------IGTKVEVVERKEHLHTDILKRGSEMDNN--CSP 388
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      Query Match
4.3%; Score 131; DB 4; Length 754;
Best Local Similarity 20.3%; Pred. No. 0.00057;
Matches 88; Conservative 69; Mismatches 161; Indels 116; Gaps
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               252 ERRAGPSRMRHERVVIIKNMFHPMDFEDDPLVLNBIR-EDLRVBCSKFGGIRKLLLFDRH 310
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              59 SECNPLLQEPIASAQFGATAGTE-CRKSVPCGWERVVKQRLFGK---TAGRFDVYFISPQ 114
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        311 -----PDGVASVSFRDPEEADYCIQTLDGRW-----FGGRQITAQAMD----- 348
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       115 GLKFRSKSSI ANYLHKNGETSLKPEDFDFTVLSKRGIKSRYKDCSMAALJSHLQNQSNNS 174
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            349 ------GTTDYQVEE----TSREREERLRGWEAFLNAPEANR 380
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            175 NWNLRTRSKCKKDVFMPPSSSSELQE--SRGLSNFTSTHLLLKEDEGVDDVNFRK---- 227
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      381 GLSVQILSLLRK---AGPSRARHFSEHPSTSKMNAQETATCMAFEEPIDEKKFEKTEDGG 437
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     13 DRGAAPTVTSSERLV-----PDPP---NDLRKEDVAMELERVGEDEEQMMIKRS 58
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                389 TRK----DFTGEKIFQEDTIPRTQIERRKTSLYFSSKYN 423
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     SEQ ID NOS: 5
FastSEQ for Windows Version 3.0
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                ; Sequence 2, Application US/09214564A; Patent No. 6150515
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                GENERAL INFORMATION:
APPLICANT: Sharp, Phillip A.
APPLICANT: Zhou, Qiang
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  ORGANISM: Homo sapiens
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              US-09-214-564A-2
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                SEQ ID NO 2
LENGTH: 754
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         US-09-214-564A-2
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               TYPE: PRT
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             SOFTWARE:
                                                                                                                                                                                                                                                                              qq
                                                                                                                                                                                                                                                                                                                                                                             δ
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   δŏ
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            δ
```

```
USA
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            FILING DATE:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               FILING DATE:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  MEDIUM TYPE:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                ADDRESSEE:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          S
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     US-08-755-587-188
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           SOFTWARE:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             COUNTRY:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         APPLICANT:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 q
                                                                                                                                                                                           QQ
                                                                                                                                                                                                                                   Qγ
                                                                                                                                                                                                                                                                                                                   δŏ
                                                                                                                                                                                                                                                                                                                                                      qq
                                                                                                                                                                                                                                                                                                                                                                                              Óζ
                                                                                                                                                                                                                                                                                                                                                                                                                                 qq
                                                                                                                                                                                                                                                                                                                                                                                                                                                                           ò
                                                                         Ωý
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 Materials and methods relating to the identification and sequencing of the BRCA2 cancer susceptibility gene and uses thereof.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           297 IGDKQAAQVSIKKDLDSSAIVHDLAEENRNSKQHLKMTLDQDLKSDISLNIXXKSDGNSD 356
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        4.2%; Score 129; DB 3; Length 1683;
20.4%; Pred. No. 0.0035;
ive 67; Mismatches 175; Indels 160; Gaps
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        11 LGDRGAAPTV----TSSERLVPDPPNDLR--KEDVAMELERVGEDEEQMMIKRSSECN-- 62
                                     438 EFFEGASENNAKESSPEKEAEEGCPEKESEEGCPK - - RGFEGSCSOKESEEGNPVRGSEE 495
                                                                             278 VAQKSQLDRTVCISDAGACGETLSVTSEENSLVKKKERSLSSGS------NFCSEQKTS 330
                                                                                                                                                          331 GIINKFCSAKDSEHNEKYEDTFLESEEIGTKVEVVER--KEHLHTDILKRGSEMDNNCSP 388
                                                                                                                                                                               545 KESDEDCSEKQSEDGSE----REFEENGLEKDLDEEGSEKELHENVLDKELEENDSENS 599
-----VRKPKGKVTILKGIPIKKTKKGCRKSCSGFVQSDSKRESVCNKADAESEP 277
                                                                                                                  496 DSPKKESKK-----KTLKNDCEENGLAKESEDDLNKESEEEVGPTKESEEDDSE
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         3: Bell Seltzer Park & Gibson
310 UCB Plaza, 3605 Glenwood Avenue, PO Drawer 31107
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                Patentin Release #1.0, Version #1.25 (EPO)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              APPLICATION NUMBER: GB 9523959.6
FILING DATE: 23-NOV-1995
PRIOR APPLICATION DATA:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 APPLICATION NUMBER: GB 9617961.9
FILING DATE: 28-AUG-1996
ATTORNEY/AGENT INFORMATION:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        UMBER: US/08/755,587
25-NOV-1996
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           UMBER: GB 9525555.0
14-DEC-1995
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             COMPUTER: IBM PC compatible OPERATING SYSTEM: PC-DOS/MS-DOS
                                                                                                                                                                                                                                                                                                                                                                               Sequence 183, Application US/08755587 Patent No. 6045997
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     APPLICANT: Ashworth, Alan
APPLICANT: Stratton, Michael R
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                RECISTRATION NUMBER: 31,665
REFERENCE/DOCKET NUMBER: 540
INFORMATION FOR SEQ ID NO: 183:
                                                                                                                                                                                                                                                                                                                                                                                                                           GENERAL INFORMATION:
APPLICANT: Futreal, Phillip A
APPLICANT: Wooster, Richard F
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           SEQUENCE CHARACTERISTICS:
LENGTH: 1683 amino acids
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     Query Match 4.2%;
Best Local Similarity 20.4%;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          MEDIUM TYPE: Floppy disk
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        Matches 103; Conservative
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    Kenneth D Sibley
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       CURRENT APPLICATION DATA:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                PRIOR APPLICATION DATA:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  PRIOR APPLICATION DATA:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           COMPUTER READABLE FORM:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        CORRESPONDENCE ADDRESS:
                                                                                                                                                                                                                                              389 TRKDFTGEKIFQED 402
                                                                                                                                                                                                                                                                                       600 EFEDDGSEKVLDEE 613
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              APPLICATION NUMBER:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               APPLICATION NUMBER:
FILING DATE: 14-DEC
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        NUMBER OF SEQUENCES:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     amino acid
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          linear
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            TITLE OF INVENTION:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           TITLE OF INVENTION:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    STREET: 310
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           USA
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 FILING DATE:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        US-08-755-587-183
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               ADDRESSEE:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        TOPOLOGY:
                                                                                                                                                                                                                                                                                                                                                                   US-08-755-587-183
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           COUNTRY:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        STATE:
            228 ---
                                                                                                                                                                                                            q
                                                                                                                                                                                                                                                    ŏ
                                                                                                                                                                     δλ
                                                                                         ŏ
                                                                                                                             Op
```

```
identification and sequencing of the BRCA2 cancer susceptibility gene and uses thereof.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               285 DRTVCISDAGACGETLSVTSEENSLVKKKERSLSSGSNFCSEQKTSGIINKFCSAKDSEH 344
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             612 -----ESDIENISEETSAEVDPISLSSSKYHD-SVASMFKIENQNXD------KSSEK 657
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                345 NEKYEDTFLESEEIGTKVEVVERVERKEHLHTDILKRGSEMDNN-----CSPTRKDFTGEK 397
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                658 NNKCQLILQNNIEMITGIEVEE----NTENYKRNTENEDNKYTGASRNSXNLEXSDGSD 712
                                                                                                                                                                                                                                                                                                                                                                                                               240 -----GIPIK----KTKKGCRKSCSGFVQSDSKRE-----SVCNKADAESEPVAQKSQL 284
                                                                                                                                                                                                                                                                                                                                                                                                                                                            552 QFEGSAGIKQKFACLLKSSCNKSASGYLTDENEVEFRGFYSALGTKLNVSSEALQKAVKL 611
                                                                                                                                                                                                        1.70 ------QSNNSNWNLRTRSKCKKDVFMPPSSSSELQESRGLSNFTSTHLLLKED 21.7
                                                                                                                                                                                                                                                      455 DSHTAPQMLSSKQDFNSNHNL------TPSQKAEITE-----LSTIL----E 491
                                                                                                                                                                                                                                                                                                                                                            492 ESGSQFEFTQFRKPSHIAQKNTFEVPENQMTILNTTSEEWKDADLHLIVNAPSIQVDSSK 551
63 -----PLLQEP1ASAQFGATAGTECRKSVPCGWERVVKQRLFGKTAGRFDVYFISPQGLK 117
                                                  357 YMDKWAGLLDPISNHSFGGSFRTASNKEIKLSEHNIKKSKMFFKDIE--EQYPTSLACVE 414
                                                                                                   118 FRSKSSLANYLHKNGETSLKPEDFDFTVLSKRGIKSRYKDCSMAALTSHLQN----- 169
                                                                                                                                                      : Bell Seltzer Park & Gibson
310 UCB Plaza, 3605 Glenwood Avenue, PO Drawer 31107
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            Materials and methods relating to the
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      PatentIn Release #1.0, Version #1.25 (EPO)
                                                                                                                                                                                                                                                                                                                 218 EGVDDVNFRKVRK-------PKGKVTILK-------
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             REFERENCE/DOCKET NUMBER: 5405-135
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      IMBER: US/08/755,587
25-NOV-1996
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         GB 9525555.0
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               GB 9617961.9
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    GB 9523959.6
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      COMPUTER: IBM PC compatible OPERATING SYSTEM: PC-DOS/MS-DOS
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           398 IFQEDTIPRTQIERRKTSLYFSSKY 422
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    Sequence 188, Application US/08755587 Patent No. 6045997
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              713 SSKNDTV---YIHKDETDLPFIDQH 734
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 APPLICANT: Ashworth, Alan
APPLICANT: Stratton, Michael R
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 Phillip A
Richard F
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          31,665
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            FILING DATE: 14-DEC 1995
PRIOR APPLICATION DATA:
APPLICATION NUMBER: GB 96
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  ATTORNEY/AGENT INFORMATION:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             28-AUG-1996
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            23-NOV-1995
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   Floppy disk
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     Kenneth D Sibley
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 CURRENT APPLICATION DATA:
APPLICATION NUMBER: US
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    FILING DATE: 25-NOV-:
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 0
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 PRIOR APPLICATION DATA:
APPLICATION NUMBER:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     REGISTRATION NUMBER:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     CORRESPONDENCE ADDRESS:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        COMPUTER READABLE FORM:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    APPLICANT: Futreal, APPLICANT: Wooster,
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               NUMBER OF SEQUENCES:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  TITLE OF INVENTION:
TITLE OF INVENTION:
TITLE OF INVENTION:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              GENERAL INFORMATION:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     Raleigh
```

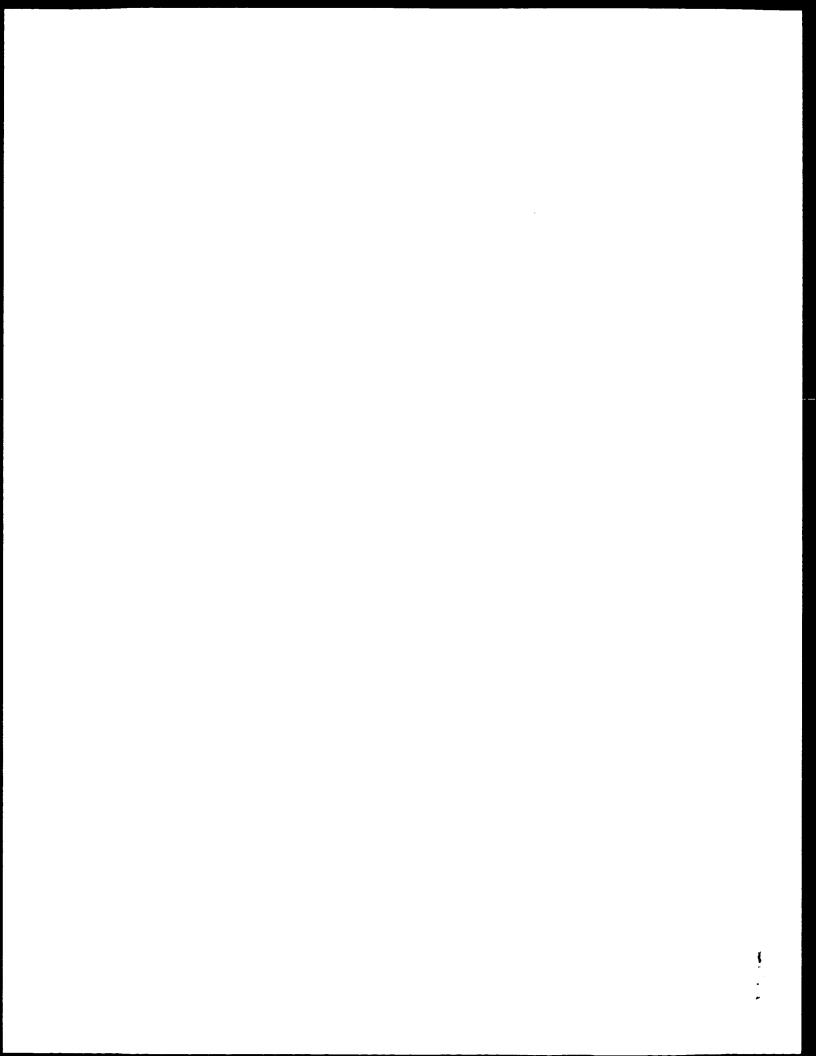
```
qq
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     δy
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    q
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         q
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  ōλ
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     qq
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           δ
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         ò
                                                                                                                                                                                    15;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                         -----GKVTILKGIPIKKTKKGCRK-----SCSGFVQSDSKRESVCNKADAESEPV 278
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          279 AQKSQLDRTVCISDA------GACGETLSVTSE-------ENSLVK 311
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              176 WNLRTRSKCKKDVFMPPSSSSELQESRGLSNFTSTHLLLKEDEGVDDVNFRKVRKPK--- 232
                                                                                                                                                                                                                                                                                                                                                                                                            407 HNL------TPSQKAEITE-----LSTIL----EESGSQFEFTQFKKPSHVA 443
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    312 KKERSLSSGSNFCSEQKTSGI----INKFCSAKDSEHNEKYEDTFLESEEIGTKVEVVER 367
                                                                                                                                             Query Match 3.8%; Score 117; DB 3; Length 1471;
Best Local Similarity 22.4%; Pred. No. 0.043;
Matches 85; Conservative 45; Mismatches 148; Indels 102; Gaps
                                                                                                                                                                                                                       63 PLLQEPIASAQFGATAGTECRKSVPCGWERVVKQRLFGKTAGRFDVYFISPQGLKFRSKS 122
                                                                                                                                                                                                                                                           297 PGFLDPVFKHKFGGSFRTASNKEIKLSEHNVKKSKMFFKDIE--EEYPTSLTCIDIVNAS 354
                                                                                                                                                                                                                                                                                                123 SLANYLHKNGETSLKPEDFDFTVL-----SKRGIKSRYKDCSMAALTSHLQNQSNNSN 175
                                                                                                                                                                                                                                                                                                                                   355 PLANQEILSG----PYTFDLQVTTMSAHPQSQASVSCEDTHTSLQVLPS---KQDFHSN 406
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            TITLE OF INVENTION: ORIGIN OF REPLICATION COMPLEX GENES NUMBER OF SEQUENCES: 24 CORRESPONDENCE ADDRESS:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 : FLEHR, HOHBACH, TEST, ALBRITTON & HERBERT 4 Embarcadero Center, Suite 3400
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.30
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               APPLICATION NUMBER: US/08/484,105
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           Sequence 14, Application US/08484105 Patent No. 5589341
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              MCNALLY, Francis J
LAURENSON, Patricia
INFORMATION FOR SEQ ID NO: 188:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               368 KEHLHTDILKRGSEMDNNCS 387
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   |: ::| |
615 PEN----YARNTKCEDNNS 629
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       BELL, Stephen P
KOBAYASHI, Ryuji
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     HERSKOWITZ, Ira
LI, Joachim J
GAVIN, Kimberly
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     APPLICANT: STILLMAN, Bruce
               SEQUENCE CHARACTERISTICS:
LENGTH: 1471 amino acids
TYPE: amino acid
TOPOLOGY: linear
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            APPLICANT: RINE, Jasper
APPLICANT: FOSS, Margit
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           CURRENT APPLICATION DATA:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               STREET: 4 EMBALCOCITY: San Francisco CTATE: California
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             94111-4187
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    GENERAL INFORMATION:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              USA
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  FILING DATE:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     ADDRESSEE:
                                                                                            US-08-755-587-188
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             US-08-484-105-14
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       APPLICANT:
APPLICANT:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              APPLICANT:
APPLICANT:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            COUNTRY:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       APPLICANT:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   APPLICANT:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         APPLICANT:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            APPLICANT:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         RESULT 13
                                                                                                                                                                                                                                                                                                                                                                                                                                                     233
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      ŏ
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               Óλ
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                Q
                                                                                                                                                                                                                                                                                                                                                                             ò
                                                                                                                                                                                                                                                                                                                                                                                                              q
                                                                                                                                                                                                                                                                                                                                                                                                                                                   ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              δ
                                                                                                                                                                                                                             δ
                                                                                                                                                                                                                                                                                                    ò
```

```
397 LEERFRTVSAKGKMETIFSKVKKQLNSRNSKEEIVKAADFDNYLPARENEFASIYLSLYS 456
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    278 VAQKSQLDRTVCISDAGACGETLSVTSEE-----NSLVKKKERSLSSGSNFCSEQKTSGI 332
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     333 INKFCSAKDSEHNEKYEDTFLESEEIGTKVEVVERKEHLHTDILKRGSEMDNNCSPTRKD 392
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      291 -EEFANASSSDSDEEFED-YQSAEELAIVEPAKKKVRSIKPDIPISPVKSQTPLQPSAVH 348
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    393 FTGEKIFQEDTIPRTQIERRKTSLYFSSKYNKEALSPPRRKAFKKWTPPRSPFNLVQETL 452
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    349 SSPRKFFKNNI-----VRAKKAYTPFSKRY-KNPKIPDLNDIFQRHN-----NDLDIAA 396
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 490 -------VARTADWRDVSELLKPLGLYDLRAKTIVKFSDEYLTKQWKYPI 532
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               457 AIEAGTSTSIYIAGTPGVGKTLTVREV------VKDLMTSADQKELPRFQY-I 502
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        222 DVNF--RKVRKPKGKVTI--LKGIPIKKTKKGCRKSCSGFVQSDSKRESVCNKADAESEP 277
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        65; Mismatches 141; Indels 103; Gaps
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  453 FHDPWKLLIA----TIF----LNRTSGKMAIPVLWKFLEKYPSAE-------
                                                                                                                                                                                                                                                                                                                                                                                                                                                  Query Match
3.8%; Score 116.5; DB 1; Length 885;
Best Local Similarity 20.4%; Pred. No. 0.02;
Matches 79; Conservative 65; Mismatches 141; Indels 103.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         GAVIN, Kimberly
JENTION: ORIGIN OF REPLICATION COMPLEX GENES
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  E: FLEHR, HOHBACH, TEST, ALBRITTON & HERBERT 4 Embarcadero Center, Suite 3400
                                                                                      REFERENCE/DOCKET NUMBER: A-59032/DJB/RAO TELECOMMUNICATION INFORMATION:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    533 ELHGIG-KYGNDSYRIFCVNEWKQVHPE 559
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      ATTORNEY/AGENT INFORMATION:
NAME: Osman Ph.D., Richard Aron
REGISTRATION NUMBER: 36,627
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             Sequence 14, Application US/08484106 Patent No. 5614618
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            RINE, Jasper
FOSS, Margit
MCNALLY, Francis J
LAURENSON, Patricia
                                                                                                                                  TELEPHONE: (415) 494-870
TELEFAX: (415) 494-8771
TELEX: 910 277299
INFORMATION FOR SEQ ID NO: 14:
SEQUENCE CHARACTERISTICS:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              BELL, Stephen P
KOBAYASHI, Ryuji
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           HERSKOWITZ, Ira
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     STILLMAN, Bruce
                                                                                                                                                                                                                                                               885 amino acids
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    LI, Joachim J
                                                                                                                                                                                                                                                                                                                                    ; TOPOLOGY: linear; MOLECULE TYPE: protein US-08-484-105-14
                                                                                                                                                                                                                                                                                                                single
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        CORRESPONDENCE ADDRESS
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             CITY: San Francisco
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           APPLICANT: GAVIN, KİR
TITLE OF INVENTION: (
NUMBER OF SEQUENCES:
                                                                                                                                                                                                                                                                                         TYPE: amino acid
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              GENERAL INFORMATION:
APPLICANT: STILLM
CLASSIFICATION:
                                                                                                                                                                                                                                                                                                                STRANDEDNESS:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          ADDRESSEE:
STREET: 4
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   APPLICANT:
APPLICANT:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      US-08-484-106-14
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              APPLICANT:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        APPLICANT:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            APPLICANT:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    APPLICANT:
APPLICANT:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    APPLICANT:
                                                                                                                                                                                                                                                                  LENGTH:
```

```
identification and sequencing of the BRCA2 cancer susceptibility gene and uses thereof.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        644 NKYTAASRNSRNLEFVGSDSSKNDTVCIHKDEKDLPFT--DQRNICLKLSGQLMKE 697
                                                                                                                                                                                                                                                                                                                                                                                                                                              Patentin Release #1.0, Version #1.25 (EPO)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            JMBER: US/08/755,587
25-NOV-1996
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              APPLICATION NUMBER: GB 9523959.6
FILING DATE: 23-NOV-1995
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    GB 9525555.0
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    GB 9617961.9
                                                                                                                                                                                                                                                                                                                                                                                                         OPERATING SYSTEM: PC-DOS/MS-DOS
                                                                                                                                                                                                                                                                                                                                           E: Floppy disk
IBM PC compatible
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  31,665
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            INFORMATION FOR SEQ 1D NO: 185;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           PRIOR APPLICATION DATA:
APPLICATION NUMBER: GB 96
FILING DATE: 28-AUG-1996
ATTORNEY/AGENT INFORMATION:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               14-DEC-1995
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           SEQUENCE CHARACTERISTICS:
LENGTH: 1535 amino acids
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 REFERENCE/DOCKET NUMBER:
                                                                                     222
                                                                                                                                                                                                                                                                                                                                                                                                                                                                   CURRENT APPLICATION DATA:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                Kenneth D Sibley
                                                                                  NUMBER OF SEQUENCES: 2:
CORRESPONDENCE ADDRESS:
                                                                                                                                                                                                                                                                                                            COMPUTER READABLE FORM: MEDIUM TYPE: Floppy
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  PRIOR APPLICATION DATA: APPLICATION NUMBER: (
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      PRIOR APPLICATION DATA;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           REGISTRATION NUMBER:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        APPLICATION NUMBER:
FILING DATE: 25-NOV
                 TITLE OF INVENTION:
TITLE OF INVENTION:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            amino acid
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  linear
                                                                                                                                                                                                             Raleigh
                                                                                                                                                                                                                                                                                 USA
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   FILING DATE:
                                                                                                                                               ADDRESSEE:
                                                                                                                                                                                                                                                      Š
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      US-08-755-587-185
                                                                                                                                                                                                                                                                                                                                                                                 COMPUTER:
                                                                                                                                                                                                                                                                                                                                                                                                                                              SOFTWARE:
                                                                                                                                                                                                                                                                                 COUNTRY:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           TOPOLOGY:
                                                                                                                                                                                     STREET:
                                                                                                                                                                                                                                               STATE:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    ŏλ
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           qq
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        qq
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   qq
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  δλ
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      18;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              278 VAQKSQLDRTVCISDAGACGETLSVTSEE----NSLVKKKERSLSSGSNFCSEQKTSGI 332
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               236 VATKLQSDNT---PSKKSFQQTVSKTNAEVQRIASTIVNEKE-AISDNESDLSEYHESK- 290
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  :| :| |: | :: || : :|| : :|| : :|| : :|| |: || |: || |: || |: || |: || |: || |: || |: || |: || |: || |: || |: || |: || |: || |: || |: || |: || |: || |: || |: || |: || |: || |: || |: || |: || |: || |: || |: || |: || |: || |: || |: || |: || |: || |: || |: || |: || |: || |: || |: || |: || |: || |: || |: || |: || |: || |: || |: || |: || |: || |: || |: || |: || |: || |: || |: || |: || |: || |: || |: || |: || |: || |: || |: || |: || |: || |: || |: || |: || |: || |: || |: || |: || |: || |: || |: || |: || |: || |: || |: || |: || |: || |: || |: || |: || |: || |: || |: || |: || |: || |: || |: || |: || |: || |: || |: || |: || |: || |: || |: || |: || |: || |: || |: || |: || |: || |: || |: || |: || |: || |: || |: || |: || |: || |: || |: || |: || |: || |: || |: || |: || |: || |: || |: || |: || |: || |: || |: || |: || |: || |: || |: || |: || |: || |: || |: || |: || |: || |: || |: || |: || |: || |: || |: || |: || |: || |: || |: || |: || |: || |: || |: || |: || |: || |: || |: || |: || |: || |: || |: || |: || |: || |: || |: || |: || |: || |: || |: || |: || |: || |: || |: || |: || |: || |: || |: || |: || |: || |: || |: || |: || |: || |: || |: || |: || |: || |: || |: || |: || |: || |: || |: || |: || |: || |: || |: || |: || |: || |: || |: || |: || |: || |: || |: || |: || |: || |: || |: || |: || |: || |: || |: || |: || |: || |: || |: || |: || |: || |: || |: || |: || |: || |: || |: || |: || |: || |: || |: || |: || |: || |: || |: || |: || |: || |: || |: || |: || |: || |: || |: || |: || |: || |: || |: || |: || |: || |: || |: || |: || |: || |: || |: || |: || |: || |: || |: || |: || |: || |: || |: || |: || |: || |: || |: || |: || |: || |: || |: || |: || |: || |: || |: || |: || |: || |: || |: || |: || |: || |: || |: || |: || |: || |: || |: || |: || |: || |: || |: || |: || |: || |: || |: || |: || |: || |: || |: || |: || |: || |: || |: || |: || |: || |: || |: || |: || |: || |: || |: || |: || |: || |: || |: || |: || |: || |: || |: || |: || |: || |: || |: || |: || |: || |: || |: || |: || |: || |: || |: || |: || |: || |: || |: || |: || |: || |
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                333 INKFCSAKDSEHNEKYEDTFLESEEIGTKVEVVERKEHLHTDILKRGSEMDNNCSPTRKD 392
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           393 FTGEKIFQEDTIPRTQIERRKTSLYFSSKYNKEALSPPRRKAFKKWTPPRSPFNLVQETL 452
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     349 SSPRKFFKNNI-----VRAKKAYTPFSKRY-KNPKIPDLNDIFORHN-----NDLDIAA 396
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         453 FHDPWKLLIA-----TIF-----LNRTSGKMAIPVLWKFLEKYPSAE-------- 489
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  ------VARTADWRDVSELLKPLGLYDLRAKTIVKFSDEYLTKQWKYPI 532
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     457 AIEAGTSTSIYIAGTPGVGKTLTVREV------VKDLMTSADQKELPRFQY-I 502
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   Sequence 185, Application US/08755587
Patent No. 6045997
GENERAL INFORMATION:
APPLICANT: Futreal, Phillip A
APPLICANT: Moster, Richard F
APPLICANT: Ashworth, Alan
APPLICANT: Stratton, Michael R
TILE OF INVENTION: Materials and methods relating to the
                                                                                                                                                                                                                                 PatentIn Release #1.0, Version #1.30
                                                                                                                                                                                                                                                                                                                                                                                    ATTORNEY/AGENT INFORMATION:
NAME: Osman Ph.D., Richard Aron
RECISTRATION NUMBER: 36,627
REFERENCE/DOCKET NUMBER: A-59032/DJB/RAO
TELECOMUNICATION INFORMATION:
TELEFRONE: (415) 494-8770
TELEFRAX: (415) 494-8771
INFORMATION FOR SEQ ID NO: 14:
SEQUENCE CHARACTERISTICS:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   533 ELHGIG-KYGNDSYRIFCVNEWKQVHPE 559
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            APPLICATION NUMBER: US/08/484,106
                                                                                                                                                                                                      PC-DOS/MS-DOS
                                                                                                                                                                IBM PC compatible
                                                                                                                                     Floppy disk
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              LENGTH: 885 amino acids
TYPE: amino acid
STRANDEDNESS: single
                                                                                                                                                                                                                                                            CURRENT APPLICATION DATA
                                                                                            COMPUTER READABLE FORM: MEDIUM TYPE: Floppy of
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              ; TOPOLOGY: linear
; MOLECULE TYPE: protein
US-08-484-106-14
                                                                                                                                                                                                                                                                                                                                                            CLASSIFICATION: 435
California
                                                                                                                                                                                                  OPERATING SYSTEM:
                                                                94111-4187
                                         USA
                                                                                                                                                                                                                                                                                                                             FILING DATE:
                                                                                                                                                                    COMPUTER:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  US-08-755-587-185
                                  COUNTRY:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        490
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                RESULT 15
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               qq
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        qq
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                δλ
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           δλ
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   qq
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        δ
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             pp
```

```
26;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       Query Match 3.8%; Score 116.5; DB 3; Length 1535; Best Local Similarity 20.4%; Pred. No. 0.051; Matches 97; Conservative 75; Mismatches 163; Indels 141; Gaps
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      293 MTLGQDLQPD------ILNIDKIPDK IDDCMDKWA---GPL--DPISNHSFGGSFR 337
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    338 TASNKEIKVSEHNIKKSKMFFK----DIEEQYPTSL---ACVEIVNTLALDNOKKLSKP 389
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      198 LQESRGLSNFTSTHLLLKEDEGVDDVNFRKVRK-------PKGKVTILK---- 239
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       80 TECRKSVPCGWERVVKQRLFGKTAGRFDVYFISPQGLKFRSKSSLANYLHKNGETSLKPE 139
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             140 DFDFTVLSKRGIKSRYKDCSMAALTSHL--QNQSNNSNWNLRTRSKCKKDVFMPPSSSSE 197
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    390 QIN-TVSAHLQSSVVVSDCKNSHITPQMSFSKQDFNSNHNL-----TPSQKAE 436
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 437 ITE-----LSTIL----EESGSQFEFTQFRKPSYILQKNTFEVPENQVTILNTTTEEN 485
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        240 ------GIPIKKTKK-----GCRKSCSGFVQSDSKRESVCNKADAESEPVAQKSQ 283
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 486 RDAGLVIMNAPSIQVNSSKQFEGTVGIKQKFAGLLKSD-----CNK--SASGYLTDENE 537
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          284 LDRTVCISDAGACGETLSVTSEENSLVKKKERSLSSGSNFCSEQKTSGII------NKF 336
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                538 VEFRGFYS---AHGVKLNVSTE---ALQKAVKLFSDIENI--SEKTSAEVDPISLSSSKF 589
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            337 CSA-----KDSEHNEKYED-----TFLESE---EIGTKVEVVERKEHLHTDILKRGSEMD 383
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  590 HDSVVSMFKIENHNDKTSEKNKCQLMCQNNIEMTTGTFVEEITENYKINTE-----NED 643
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         20 VTSSERLVPDPPNDLRKEDVAMELERVGEDFEQMMIKRSSECNPLLQEPIASAQFGATAG 79
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         384 NNCSPTRK-----DFTGEKIFQEDTI-----PRTQIERRKTSLYFSSKYNKE 425
E: Bell Seltzer Park & Gibson
310 UCB Plaza, 3605 Glenwood Avenue, PO Drawer 31107
```

Search completed: January 15, 2002, 13:51:25 Job time: 354 sec



```
GenCore version 4.5 Copyright (c) 1993 - 2000 Compugen Ltd.
```

OM protein - protein search, using sw model

January 15, 2002, 13:52:32 Run on:

; Search time 74.94 Seconds
(without alignments)
589.555 Million cell updates/sec

Title: Perfect score:

US-09-724-296-38 3055 1 MGTTGLESLSLGDRGAAPTV.......HKLNKYHDWLWENHEKLSLS 580 Sequence:

BLOSUM62 Scoring table:

219241 seqs, 76174552 residues Gapop 10.0 , Gapext 0.5 Searched:

219241 Total number of hits satisfying chosen parameters:

Minimum DB seq length: 0 Maximum DB seq length: 200000000

Post-processing: Minimum Match 0% Maximum Match 100% Listing first 45 summaries

Database :

PIR_68:*
1: pir1:*
2: pir2:*
3: pir3:* 4: pir4:* Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

rhoptry-associated nucleolar protein protein YKR029c ho cylicin I - human hypothetical prote probable ABC trans protein kinase (EC nebulin, skeletal hypothetical prote ankyrin 2, neurona hypothetical prote cag pathogenicity hypothetical prote hypothetical prote unknown protein [i methyl CpG binding methyl-CpG-binding SAC3 protein - yea breast cancer tumo cag island protein cut17 protein - fi hypothetical prote hypothetical prote hypothetical prote activator 1 large hypothetical prote probable A/G-speci DNA-binding protei Description SUMMARIES S57963 A41907 S51323 G02334 S27833 S55785 S55785 T34036 T41622 A72682 A71928 T43523 G64585 H86442 T50346 T33350 T14050 S55024 C86333 S37431 S51441 JN0599 A49393 T33813 S55101 Query Match Length DB 4476 13091 13091 13091 1359 1359 1359 1927 1927 1927 1002 1000 1000 1000 209 209 145 141.5 1341.5 137.5 136.5 136.5 135.5 131.5 131.5 131.5 128.5 Score Result

: | |||::: 345 ---SSSASSPPKKE 355

qq

hypothetical prote

nuclear/mitotic ap	hypothetical prote	DNA-directed DNA p	hypothetical prote	hypothetical prote		major surface qlvc	hypothetical prote	hypothetical prote		nucleolus-cytoplas	neurofilament prot	transcription requ	breast cancer susc	breast cancer tumo	hypothetical prote
T30336	T43773	T17202	F71620	H84919	н86168	T30546	B84614	E71412	B84590	B42680	A46194	T51447	T42205	T30904	T10633
7	7	7	7	~	a	7	7	7	7	7	7	7	C*	~	7
53	1740	3122	646	797	166	1002	891	1040	533	734	1200	. 1377	3329	3329	957
22															
3.9 22	3.9	9. 6.	9. g	3.9	9.8	3.9	3.9	3.9	3.9	3.9	3.9	3.9	3.9	3.9	3.8
3.9	119.5 3.9 1	119.5 3.9	119 3.9	119 3.9	119 3.9	119 3.9	118.5 3.9	118.5 3.9	118 3.9	118 3.9	e 	118 3.9	118 3.9	118 3.9	117.5 3.8

ALIGNMENTS

RESULT 1 857963 methyl CpG binding protein 2 - human (fragment) C.Species: Homo sapiens (man) C.Date: 19-Mar-1997 #sequence_revision 25-Apr-1997 #text_change 05-Nov-1999	C:Accession: S5/7963 R:A'Esposito, M.: Quaderi, N.A.; Ciccodicola, A.; Bruni, P.; Esposito, T.; D'Urso, submitted to the EMBL Data Library, July 1995	A;Description: Physical mapping and expression analysis of an X-linked gene encoding A;Reference number: S57963	A,Accession: 55/963 A,Status: preliminary	A:Molecule type: mRNA A:Residues: 1-476 <pre>chrs></pre>	A:Cross-references: EMBL:X89430; NID:9899295; PIDN:CAA61599.1; PID:9899296	Query Match 6.8%; Score 209; DB 2; Length 476; Best Local Similarity 24.0%; Pred. No. 4.1e-06; Matches 104; Conservative 53; Mismatches 149; Indels 128; Gaps 17;	37 EDVAMELERVGEDEEQMMIKRSSECNPLLQEPIASAQFGATAGTE 81	11 :: :: : : : : : : : : : : :	82 CRKSVPCGWERVVKQRLFGKTAGRFDVYFISPQGLKFRSKSSLANYL 128	73 QRRSIIRDRGPMYDDPTLPEGWTRKLKQRKSGRSAGKYDVYLINPQGKAFRSKVELIAPF 132	129 HKNGETSLKPEDFDFTVLSKRGIKSRYKDCSMAALTSHLQNQSNNSNWHLKTRSKCKKDV 188	133 EKVGDTSLDPNDFDFTV-TGRGSPSR	189 EMPPSSSSELQESRGLSNFTSTHLLLKEDEGVDDVNFRKVRRPKGKVTILKGIPIKK 245	166PKSPKAPGTGRGRPKGSGTTRPKAATSEGVQVKRVLEKSPGKLLVKMPF-Q 217	246 TKKGCRKSCSGFVQSDSKRESVCNKADAESEPVAQKSQLDRTVCISDAGACGETLS 301	18 TŚPGGKAEGGGATTSTQVMVIKRPGRKRKAEADPQAIPKKRGRKPGSVVA 267	102 VTSEENSLVKKKERSLSSGSNFCSEQKTSGIINKFCSAKDSEHNEKYEDTFLESEEIGTK 361	268 AAAAEAKKKAVKESSIRVS 303	362 VEVVERKEHLHTDILKRGSEMDNNCSPTRKDFTGEKIFQEDTIPRTQIERRKTSLYFS 419	304 IEVKEVVKPLLVSTLGEKSGKGLKTCKSPGRKSKESSPKGR 344	420 SKYNKEALSPPRRK 433
inc lomc lar-	n: S: to, b; to t;	ion: e num	on: s: preli	: type	ferer	itch al Si 104;	EDVAN	KDKPI	CRKS\	QRRSI	HKNGE	EKVGE	FMPPS	PK	TKKGC	TSPGC	VTSEE	AAAAE	VEVVE	IEVKE	SKYNK
77.7.5		pt nc.	ĭ	ule	-re	Ma Loc	37	13	82	73	129	133	189	166	246	218	0.5	898	62	304	20
RESULT 1 S57963 methyl CpG b C;Species: H	essio sposi tted	cri	tus	0 -0	S	ry t J									• •		(+)	1.4	(7)	.,	4

```
A; Description: The SAC3 gene codes for a nuclear protein required for normal mitosis.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    A; Cross-references: EMBL: Z47805; NID: g634085; PIDN: CAA87767.1; PID: g634086 C; Genetics:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          629 L-----STNPLLTPQVHGDLSEQKQQQIKTVTDG----GSPFVFDQSAQNSTVEASK 676
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              Best Local Similarity 21.9%; Pred. No. 0.22;
Matches 116; Conservative 61; Mismatches 185; Indels 168; Gaps
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             510 TLQHYSHKLSETQPLKKTY-LTCLERRLQKTTYKGLINGGEDNLASSVYVKDPKKDRIPS 568
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                160 MAALTSHLONOSNNSNWNLRTRSKCK-----KDVFMPPSSSSELQESRGLSNFTSTHL 212
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   569 IADQSFLMENFONNYNEKLNONSSVKPQINTSPKRVATRPNHFPFSQESKQLPQISQSHT 628
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      213 LLKEDEGVDDVNFRKVRKPKGKVTILKGIPIKKTKKGCRKSCSGFVQSDSKRESVCNKAD 272
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                273 AESEPVAÇKSÇLDRTVCISDAGACGETLSVTSEENSLVKKKERSLSSGSNFCSEQKTSGI 332
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   677 A····---HMISTTSNGAYDEKLS·-SEQEEMRKKEEQRI------EEEKT--- 712
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      333 INKFCSAKDSEHNEKYEDTFLESEEIGTKV--EVVERKEHLHTDILKRGSEMDNNCSPTR 390
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             391 KDFTGEK1FQEDT1PR----TQ1ERRKTSLYFSSKYNKEALSPPRRKAFKKW------ 438
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            758 KDFI-----DTMTRELYDAFLHERLYLIYMDSRAELKRNSTLKKKFFEKWQASYSQAK 810
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   439 ----TFPRSPFNLVQE----TLFHDPWKLLIATIFLNRTSGKMAI--PV-- 477
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      811 KNRILEEKKREEIKLVSHQLGVPGFKKSTCLFRTPYKGNVNSSFMLSSSDKNLIFSPVND 870
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         478 ------LWKFLE-----KYPSAEVARTADWRDVSELLKPLGLYDLR 512
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                -----SNSLTPANLF-IY 916
                                                                                                                                                                                                                                                                                                                                                                         A;Residues: 1-1301 <MUR>
A;Cross-references: EMBL;Z50046; NID:g899393; PID:g899406; MIPS:YDR159w
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  Yeast 12, 965-975, 1996
A.Title: Characterization of the SAC3 gene of Saccharomyces cerevisiae.
A.Reterence number: S71744, MLID:97027306
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          713 -----OLKKKQEN---ADKQVITEQIANDLVKEVVNSSV---ISIVKREFSEAN----YR
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     A;Description: potential regulator of leucine permease gene expression C;Keywords: nucleus; transmembrane protein F;999-1015/Domain: transmembrane #status predicted <TMM>
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                DB 2; Length 1301;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                513 AKTIVKFSDEYLTKQWKYPIELHGIGKYGND--SYRIFCVNEWKQVHPED 560
                                                                      A;Accession: S51323
A;Molecule type: DNA
A;Residues: 1-1301 <AAD
A;Cross-references: EMBL:Z47805; NID:g634085; PID:g634086
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          123 SLANYLHKNGETSLKPEDFDFTVLSKRGIKSRYKDC----
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      4.7%; Score 145; DB;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                871 EFNKFATHLTKISKLWRPLEMQSIYYDNLTKKFP----
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            A; Cross-references: SGD: S0002566; MIPS: YDR159w
                                                                                                                                                                                                         R.Murphy, L.; Richards, C.; Harris, D. submitted to the EMBL Data Library, July 1995 A; Reference number: $57971 A; Accession: $57983 A; Molecule type: DNA
                                                                                                                                                                                                                                                                                                                                                                                                                                                           A, Experimental source: strain AB972 R; Bauer, A.; Koelling, R.
                             A; Reference number: S51323
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               A. Molecule type: DNA
A. Residues: 1-1301 <BAW>
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  A; Gene: SGD: SAC3; LEP1
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     A; Map position: 4R
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       Query Match
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              C; Function:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   qq
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                δ
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       a
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          q
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                Οy
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   qq
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          qq
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                q
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          ŏ
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          pp
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 a
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       Ω
                                                                                                                                                                                                     C; Species: Ratus norvegicus (Norway ra)
C; Species: Ratus norvegicus (Norway ra)
C; Species: Ratus norvegicus (Norway ra)
C; Species: Ratus norvegicus (Norway ra)
C; Species: Ratus norvegicus (Norway ra)
C; Species: Ratus norvegicus (Norway ra)
C; Species: Ratus norvegicus (Norway ra)
C; Species: Species: Ratus Ra; Henzel, W.J.; Maurer-Fogy, I.; Jeppesen, P.; Klein, F.; Bir (Cell 69, 905-914, 1992
A; Title: Purification, sequence, and cellular localization of a novel chromosomal protein A; Reference number: A1907; MUD:92298389
A; Title: Purification, Sequence, and cellular localization of a novel chromosomal protein A; Reference number: RNA
A; Accross-references: GB:M94064; NID:9205361; PIDN:AAA41584.1; PID:9205362
R; Nan, X.; Meehan, R.R.; Bird, A.
A; Reference sees: GB:M94064; NID:94205361; PIDN:AAA41584.1; PID:9205362
R; Nan, X.; Meehan, R.R.; Bird, A.
A; Cross-references: GB:M94064; NUD:9422813
A; Title: Dissection of the methyl-CpG binding domain from the chromosomal protein McCP2.
A; Contents: annotation; methyl-CpG-binding domain
C; Keywords: chromosomal protein; DNA binding
F; 78-162/Domain: methyl-CpG-binding #status experimental <MCG>
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   18;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 N;Alternate names: protein YD8358.13; protein YDR159w
C;Species: Saccharomyces cerevisiae
C;Date: 23-Feb-1995 #sequence_revision 12-May-1995 #text_change 21-Jul-2000
C;Accession: S51323; S57983; S71744
R;Bauer, A.; Koelling, R.
submitted to the EMBL Data Library, January 1995
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           ------PCGWERVVKORLFGKTAGRFDVYFISP 113
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     68 SGSAPAVPEASASPKQRRSIIRDRGPMYDDPTLPEGWTRKLKQRKSGRSAGKYDVYLINP 127
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  114 QGLKFRSKSSLANYLHKNGETSLKPEDFDFTVLSKRGIKSRYKDCSMAALTSHLQNQSNN 173
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            174 SNWNLRTRSKCKKDVFMPPSSSSELQESRGLSNFTSTHLLLKEDEGVDDVNFRKV-RKPK 232
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        168 -----REQKPPKK-----PKSPKAPGTGRGRPKGSGTGRPKAAASEGVQVKRVLEKSP 217
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  233 GKVTILKGIPIKKTKKGCRKSCSGFVQSDS----KRESVCNKADAESEPVAQKSQLDRT 287
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         288 VCISDAGACGETLSVTSEENSLVKKKERSLSSGSNFCSEQKTSGIINKFCSAKDSEHNEK 347
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            269 -----GRKPGSVVAAAAAEAKKKAVKESSIR-----SVQETVLPIKK------RK 307
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               348 YEDTFLESEEIGTKVEVVERKEHLHTDIL--KRGSEMDNNCSPTRKDFTGEKIFQEDTIP 405
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  308 TRET-----VSIEVKEVVKPLLVSTLGEKSGKGLKTCKSPGRK-----SKESSP 351
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   Gaps
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  34 LRKE------DVAMELERVGEDEEQMMIKRSSECNPLLQEPIASAQFG----- 75
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     8 LRKEKSEDQDLQGLKEKPLKFKKVKKDKKEDKEGKHEPLQPSAHHSAEPAEAGKAETSES 67
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  218 GK -- LLVKMPFQASPGG -- KGEGGGATTSAQVMVIKRPGRKRKAEADPQAIPKKR ----
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     6.6%; Score 201; DB 2; Length 492;
22.9%; Pred. No. 1.4e-05;
Live 57; Mismatches 156; Indels 150;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     406 RTQIERRKTSLYFSSKYNKEALSPPRR-------KAFKKWTPPRSP 444
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            352 KGR-----SSSASSPPKKEHHHHHAESPKAPMPLLPPPPP 388
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       SAC3 protein - yeast (Saccharomyces cerevisiae)
                                                                                                                                             methyl-CpG-binding protein 2 - rat N;Alternate names: chromosomal protein MeCP2
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            76 ----ATAGTECRKSV---
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           Matches 108; Conservative
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              Best Local Similarity
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                Query Match
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        qq
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              QQ
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         ŏ
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           qq
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  q
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            q
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     δ
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               à
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 Q
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      òγ
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  δ
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        δ
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  ò
```

```
rhoptry-associated protein 1 precursor - malaria parasite (Plasmodium falciparum)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     535 HGIGKYGNDSYRIF-CVNEWKQVHP-----EDHKLNKYH 567
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     - CIGSFGEHHLRLYEFENDLFKFHPNIDYLTLADGYKLQKNH 451
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               A;Reference number: S55785; MUID:95319932
A;Accession: S55785
                                            protective antigen
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      Matches 110; Conservative
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              Best Local Similarity
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          A; Residues: 1-782 <RI2>
                                                                                                                                                                                                                                                                                                                                                                                                                                                                      A;Status: preliminary
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             A; Molecule type: DNA
                                                                                                                                                                                                                                                                                                                                                                                                                               A; Accession: A45514
                                                      N;Alternate names:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  Query Match
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        RESULT
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         ŏ
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           οy
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    q
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      δ
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          q
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             ô
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 qq
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    Ωý
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        9
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           δ
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    qq
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           Dp
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              δ
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       Op
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         9
                                                                                                                                                                   C;Species: Homo sapiens (man)
C;Date: 2.1-Dec-1916 #sequence_revision 06-Jun-1997 #text_change 22-Jun-1999
C;Accession: 60234; S68501
R;Tavtigian, S.V.; Rommens, J.M.; Couch, F.J.; Neuhausen, S.; Bell, R.; Berry, S.; Bogde er, M.; Snyder, S.; Stringfellow, M.; Stroup, C.; Swedlund, B.; Teng, D.; Thomas, A.; Tr submitted to the EMBL Data Library, December 1995
A;Reference number: H01078
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          A;Molecule type: mRNA
A;Residues: 282-371,'N',373-598,'S',600-1108,'EQ',1111-1119,'D',1121-2321,'V',2323-2386
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              A. Molecule type: mRNA
A. Residues: 1-3418 cravy
A. Cross-references: EMBL: U43746; NID:g1161383; PIDN: AAB07223.1; PID:g1161384
A. Cross-references: EMBL: U43746; NID:g1161383; PIDN: AAB07223.1; PID:g1161384
A. Cross-references: EMBL: U43746; NID:g1161383; PIDN: AB07223.1; PID:g1161384
B. Wooster, R.; Bignell, G.; Lancaster, J.; Swift, S.; Seal, S.; Mangion, J.; Collins, ith, A.; Connor, F.; Arason, A.; Gudmundsson, J.; Ficenec, D.; Kelsell, D.; Ford, D.; Nature 378, 789-792, 1995
A. Authors: Futreal, P.A.; Ashworth, A.; Stratton, M.R.
A. Authors: Futreal, P.A.; Ashworth, A.; Stratton, M.R.
A. Reference number: 568501; MUID: 96112016
A. Accession: 568501
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        23;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           972 KSDISLNIDKIPEKNNDYMNKWAG-----LLGPISNHSFGGSFRTASNKEIKLSEHNIKK 1026
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         1078 VSDCKNSHITPQMLFSKQDFNSNHNL------TPSQKAEITE-----LSTIL 1118
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  1027 SKMFFK----DIEEQYPTSL---ACVEIVNTLALDNQKKLSKPQSIN-TVSAHLQSSVV 1077
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                1119 ----EESGSOFEFTOFRKPSYILOKSTFEVPENOMTILKTTSEECRDADLHVIMNAPSIG 1174
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       1175 QVDSSKQFEGTVEIKRKFAGLLKNDCNKSASGYLTDE-----NEVGFRGFYSAHGTKL 1227
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 96 QRLFGKTAGRFDVYFISPQGLKFRSKSSLANYLHKNGETSL-KPEDFDFTVLSKRGIKSR 154
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                155 YKDCSMAALTSHL--QNQSNNSNWNLRTRSKCKKDVFMPPSSSSELQESRGLSNFTSTHL 212
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     285 D-RTVCISDAGACGETLSVTSEENSLVKKKERSLSSGSNFCSEQKTSGIINKFCSAKDSE 343
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                344 HNEKYEDTFLESEE------IGTKVEVVERKEHLHTDILKRGSEMDNN--CSP 388
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  36 KEDVAMELERVGEDEEQMMIKRSSECNPLLQEPIASAQFGATAGTECRKSVPCGWERVVK 95
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         1228 NVSTEALQKAVKLFSDIENISEETS---AEVHPISLSSSKCHD----SVVSMF---KIEN
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              240 ------GIPIKK-----TKKGCRKSCSGFVQSDSKRESVCNKADAESEPVAOKSQL
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 Query Match
4.7%; Score 145; DB 1; Length 3418;
Best Local Similarity 20.9%; Pred. No. 0.81;
Matches 96; Conservative 67; Mismatches 158; Indels 138;
                                                                                                                      N; Alternate names: breast cancer susceptibility protein BRCA2
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       213 LLKEDEGVDDVNFRKVRK-------PKGKVTILK------
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              :| :| | :| | 3.0 | 3.0 | 3.0 | 3.0 | 3.0 | 3.0 | 3.0 | 3.0 | 3.0 | 3.0 | 3.0 | 3.0 | 3.0 | 3.0 | 3.0 | 3.0 | 3.0 | 3.0 | 3.0 | 3.0 | 3.0 | 3.0 | 3.0 | 3.0 | 3.0 | 3.0 | 3.0 | 3.0 | 3.0 | 3.0 | 3.0 | 3.0 | 3.0 | 3.0 | 3.0 | 3.0 | 3.0 | 3.0 | 3.0 | 3.0 | 3.0 | 3.0 | 3.0 | 3.0 | 3.0 | 3.0 | 3.0 | 3.0 | 3.0 | 3.0 | 3.0 | 3.0 | 3.0 | 3.0 | 3.0 | 3.0 | 3.0 | 3.0 | 3.0 | 3.0 | 3.0 | 3.0 | 3.0 | 3.0 | 3.0 | 3.0 | 3.0 | 3.0 | 3.0 | 3.0 | 3.0 | 3.0 | 3.0 | 3.0 | 3.0 | 3.0 | 3.0 | 3.0 | 3.0 | 3.0 | 3.0 | 3.0 | 3.0 | 3.0 | 3.0 | 3.0 | 3.0 | 3.0 | 3.0 | 3.0 | 3.0 | 3.0 | 3.0 | 3.0 | 3.0 | 3.0 | 3.0 | 3.0 | 3.0 | 3.0 | 3.0 | 3.0 | 3.0 | 3.0 | 3.0 | 3.0 | 3.0 | 3.0 | 3.0 | 3.0 | 3.0 | 3.0 | 3.0 | 3.0 | 3.0 | 3.0 | 3.0 | 3.0 | 3.0 | 3.0 | 3.0 | 3.0 | 3.0 | 3.0 | 3.0 | 3.0 | 3.0 | 3.0 | 3.0 | 3.0 | 3.0 | 3.0 | 3.0 | 3.0 | 3.0 | 3.0 | 3.0 | 3.0 | 3.0 | 3.0 | 3.0 | 3.0 | 3.0 | 3.0 | 3.0 | 3.0 | 3.0 | 3.0 | 3.0 | 3.0 | 3.0 | 3.0 | 3.0 | 3.0 | 3.0 | 3.0 | 3.0 | 3.0 | 3.0 | 3.0 | 3.0 | 3.0 | 3.0 | 3.0 | 3.0 | 3.0 | 3.0 | 3.0 | 3.0 | 3.0 | 3.0 | 3.0 | 3.0 | 3.0 | 3.0 | 3.0 | 3.0 | 3.0 | 3.0 | 3.0 | 3.0 | 3.0 | 3.0 | 3.0 | 3.0 | 3.0 | 3.0 | 3.0 | 3.0 | 3.0 | 3.0 | 3.0 | 3.0 | 3.0 | 3.0 | 3.0 | 3.0 | 3.0 | 3.0 | 3.0 | 3.0 | 3.0 | 3.0 | 3.0 | 3.0 | 3.0 | 3.0 | 3.0 | 3.0 | 3.0 | 3.0 | 3.0 | 3.0 | 3.0 | 3.0 | 3.0 | 3.0 | 3.0 | 3.0 | 3.0 | 3.0 | 3.0 | 3.0 | 3.0 | 3.0 | 3.0 | 3.0 | 3.0 | 3.0 | 3.0 | 3.0 | 3.0 | 3.0 | 3.0 | 3.0 | 3.0 | 3.0 | 3.0 | 3.0 | 3.0 | 3.0 | 3.0 | 3.0 | 3.0 | 3.0 | 3.0 | 3.0 | 3.0 | 3.0 | 3.0 | 3.0 | 3.0 | 3.0 | 3.0 | 3.0 | 3.0 | 3.0 | 3.0 | 3.0 | 3.0 | 3.0 | 3.0 | 3.0 | 3.0 | 3.0 | 3.0 | 3.0 | 3.0 | 3.0 | 3.0 | 3.0 | 3.0 | 3.0 | 3.0 | 3.0 | 3.0 | 3.0 | 3.0 | 3.0 | 3.0 | 3.0 | 3.0 | 3.0 | 3.0 | 3.0 | 3.0 | 3.0 | 3.0 | 3.0 | 3.0 | 3.0 | 3.0 | 3.0 | 3.0 | 3.0 | 3.0 | 3.0 | 3.0 | 3.0 | 3.0 | 3.0 | 3.0 | 3.0 | 3.0 | 3.0 | 3.0 | 3.0 | 3.0 | 3.0 | 3.0 | 3.0 | 3.0 | 3.0 | 3.0 | 3.0 | 3.0 | 3.0 | 3.0 | 3.0 | 3.0 | 3.0 | 3.0 | 3.0 | 3.0 | 3.0 | 3.0 | 3.0 | 3.0 | 3.0 | 3.0 | 3.0 | 3.0 | 3.0 | 3.0 | 3.0 | 3.0 | 3.0 | 3.0 | 3.0 |
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  389 TRK----DFTGEKIFQEDTIPRTQIERRKTSLYFSSKYN 423
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              A:Map position: 13q12.3-13q12.3
C;Superfamily: breast cancer tumor suppressor BRCA2
                                                                               breast cancer tumor suppressor BRCA2 - human
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            A;Cross-references: GDB:387848; OMIM:600185
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            C. Keywords: polymorphism; tumor suppressor
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            A; Status: translated from GB/EMBL/DDBJ
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    A; Accession: G02334
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     A; Gene: GDB: BRCA2
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           q
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             QQ
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                õ
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       셤
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              염
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     g
```

```
Natical protein gar2 - fission yeast (Schizosaccharomyces pombe)
Nathernate names: probable RNA binding protein RBD18
Nathernate names: probable RNA binding protein RBD18
C; Decis: Schizosaccharomyces pombe
C; Date: 28-Oct-1996 #sequence_revision 07-Feb-1997 #text_change 17-Nov-2000
C; Accession: S55785; S68087; T37634; S52318
R; Gulli, M.P.; Girard, J.P.; Zabetakis, D.; Lapeyre, B.; Melese, T.; Caizerques-Ferre
Nucleic Acids Res. 23, 1912-1918, 1995
A; Title: gar2 is a nucleolar protein from Schizosaccharomyces pombe required for 188
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            25;
C;Species: Plasmodium falciparum
C;Date: 17-Apr-1993 #sequence_revision 17-Apr-1993 #text_change 09-Jun-2000
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      123 SLANYLHKNGET----SLKPEDFDF-----TVLSKRGIKSRYKDCSMAALTSHLQNQS- 171
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           172 -NNSNWNLRTRSKCKKDVFMPPSSSSELQESRGLSNFTSTHLLLKEDEGVDDVNFRKVRK 230
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        76 VDDGNINLTDTSTSNK-----SSKKGHGRSRVRSASAAAILEEDDSKDDMEF---- 122
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          231 PKGKVTILKGIPIKKTKKGCRKSCSGFVQSDSKRESVCNKADAESEPVAQKSQLDRTVCI 290
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 291 SDAGACGETLSVTSEENSLVKKKERSLSSGSNFCSEQKTSG11NKFCSAKDSEHNEKYED 350
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          159 KSASPHGES----KSSEESTTKSSKRSAS-----VAGIV-----GADEEAPPAPKN 200
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            351 TFLESEEI-GTKVEVVERKEHLHT-----DILKRGS------EMDNNCSPTRKDFTG 395
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         201 TLTPLEELYPTNVNLFNYKYSLNNMEENINILKNEGDLVAQKEEFEYDENMEKAKQDKKK 260
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               --EKI-------FQEDTIPRTQIERRKTSLYFSSKYNKEALSPPRRKAFKKWTPPRS 443
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              444 PFNLVQETLFHDPWKLLIATIFLNRTSGKMAIPVLWKFLEKYPSA------- 488
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       308 -DEVIEKT----EVSKKTFSGIGFNLTDKEAKVLGVGATYQEYPETMLYNCPNNSNLFDTI 363
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      364 ESLQGRIIDIKKRESMISTTFEQOKECLKNMGVLDLELNDTQCKFGT------ 410
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            261 ALEKIGKOSDEEPFMFSENKFLENOVKERNVAGSFSRFFSK--LNP----FKK---- 307
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 489 ------EVARTADWRDVSELLKPLGLYDLRAK-TIVKFSDEYLTKQWKYPIEL 534
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            Gaps
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 18 NVADGINVNGDNNYGKTIINNDFNFDDYNYWTPINKKEFLNSYED--EFSSESFLENKSS 75
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          71; Mismatches 176; Indels 165;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             4.6%; Score 141.5; DB 2; Length 782; 21.1%; Pred. No. 0.18;
                                                                                                                                                                                                                                                                                                                                                                                                                                                 A; Cross-references: GB:M32853; NID:g160656; PID:g160657
C; Superfamily: Plasmodium falciparum rhoptry-associated protein 1
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  A; Status: nucleic acid sequence not shown
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         A; Molecule type: DNA
A; Residues: 1-500 <GUL>
```

24;

Gaps

```
R;Rasmussen, S.W.
Yeast 11, 873-883, 1995
A;Title: A 37.5 kb region of yeast chromosome X includes the SME1, MEF2, GSH1 and CSD A;Reference number: S57357; MUID:96090136
A;Accession: S57363
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  J. Cell Biol. 122, 1043-1052, 1993
A,Title: Molecular characterization of mammalian cylicin, a basic protein of the sper
A,Reference number: A40713; MUID:93359502
                                                                                                                                                                                                                                                                                                                                                                                                                                                                A;Cross-references: EMBL:X85021; NID:g728698; PIDN:CAA59389.1; PID:g728705
A;Note: the nucleotide sequence was submitted to the EMBL Data Library, February 1995
C;Genetics:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            C;Species: Homo sapiens (man)
C;Dale: 12-May-1994 #sequence_revision 12-May-1994 #text_change 21-Jul-2000
C;Acression: B40713; S3592
S;Hess, H.; Heid, H.; Franke, W.W.
J. Cell Biol. 122, 1043-1052, 1993
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       118 FRSKSSLANYLHKNGETSLKPEDFDFTVLSKRGIKSRYKDCSMAALTSHL--QNQSNNSN 175
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    ---TLSVTSEENSLVKK-KERSLS 318
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               229 DERLEQFSSIVTTSASNTNQHQOSVNNTEEQPKKRQLHYTAPTTENSNSIRKKLRQEKLV 288
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          SGSNFCSEQKTSGIINKFCSAKD-----SEHNEKYEDTFLESEEIGTKVEVVERKEH 370
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               371 LHTDIL------KRGSEMDNNCSPTRKDFTGEKIFQEDTIPRTQ-IERRKTSLYFSSKY 422
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      1 : 1 : 1 : 1 | 330 YDDDWVVCSNWESSRSADIEVRKSSNERDF---GVFAADSCVKGELIQEYLGKIDFQKNY 386
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         423 NKEALSPPRRKAFKKWTPPRSPFNLVQET----LFHDPWKLLIATIFLNRTSGKMAIPVL 478
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    387 Q-----LTDPNNDYRLMGTTKPKVLFHPHWPLYIDS---RETGG-----L 422
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   479 WKFLEK--YPSAEVARTADWRDVSELLKPLGLYDLRAKTIVKF----SDEYLTKQWKYP 531
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             423 TRYIRRSCEPNVELVTV---RPLDE--KPRGDNDCRVKFVLRAIRDIRKGEEISVEWQWD 477
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     176 WNLRTRSKCKKDVFMPPSSSSELQESRGLSNFTSTHLLLKEDEGVDDVNFRKVRKPKGKV 235
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       135 VSV-THNESSKENKITPSMRAE-----10NKPKNGCICGSSDSKDEL------174
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           236 TILKGIPIKKTKKGCRKSCSGFVQSDS-KRESVCNKADAESEPVAQKSQLD------ 285
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  82 FONDRGIFNHHSSSGSSKT----ASTNKRGIAAAVALATAATIPFPLKKONODDNSK 134
                                                                                                                                A; Cross-references: EMBL: Z49380; NID: 91008285; PID: 91008286; MIPS: YJL105w
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 4.5%; Score 137; DB 2; Length 560;
18.7%; Pred. No. 0.23;
tive 80; Mismatches 161; Indels 150;
                                                                                                                                                                                                                                                                                                                                                    A;Status: nucleic acid sequence not shown; translation not shown A;Molecule type: DNA A;Residues: 1-560 <RAF>
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     289 VSSHFLKP----LLNEVSSSNDTEFKAITISEYKDKYVKMFIDN-----
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        286 ----RTVCISDAGACGE---
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           Best Local Similarity 18.7%
Matches 90; Conservative
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       cylicin I - human (fragment)
                                                 A; Molecule type: DNA
A; Residues: 1-560 <RAW>
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     A; Molecule type: mRNA
A; Residues: 1-598 <HES>
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    A;Status: preliminary
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          A; Accession: B40713
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     A; Map position: 10L
            A; Accession: S56883
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 532 I 532
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            478 L 478
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              Query Match
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    ω
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    RESULT
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                qq
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   δy
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    QΫ
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   qq
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  δλ
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     a
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            qq
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        q
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   ŏ
A;Cross-references: EMBL:248166; NID:g663261; PIDN:CAA88179.1; PID:g663262
R;VanHoy, R.W.; Wise, J.A.
Cutr. Genet. 29, 307-315, 1996
A;Title: Molecular analysis of a novel Schizosaccharomyces pombe gene containing two RNF
A;Reference number: S68083; MUID:96171513
A;Reference number: S68087
A;Status: nucleic acid sequence not shown; not compared with conceptual translation
A;Retacus: nucleic acid sequence not shown; not compared with conceptual translation
A;Retacus: nucleic acid sequence not shown; not compared with conceptual translation
A;Retacus: nucleic acid sequence not shown; not compared with conceptual translation
A;Retacus: nucleic acid sequence not shown; not compared with conceptual translation
A;Retacus: D: Churcher C.M.; Barrell, B.G.; Rajandream, M.A.; Wood, V.
R;Brown, D.; Churcher, C.M.; Barrell, B.G.; Rajandream, M.A.; Wood, V.
R;Brown, D.; Churcher, C.M.; Barrell, B.G.; Rajandream, M.A.; Wood, V.
A;Residues: 1-37634
A;Status: preliminary; translated from GB/EMBL/DDBJ
A;Accession: T37634
A;Status: preliminary; translated from GB/EMBL/DDBJ
A;Accession: T37634
A;Status: preliminary; translated from GB/EMBL/DDBJ
A;Accession: T37634
A;Residues: 1-338,'S', 340-500 cBRO>
A;Cross-references: EMBL:299091; PIDN:CABI1772.1; GSPDB:GN0066; SPDB:SPAC13F5.09
A;Genetics:
A;Genetics:
A;Genetics:
A;Genetics:
A;Genetics:
A;Genetics:
A;Genetics:
A;Genetics:
A;Genetics:
A;Genetics:
A;Genetics:
A;Genetics:
A;Genetics:
A;Genetics:
A;Genetics:
A;Genetics:
A;Genetics:
A;Genetics:
A;Genetics:
A;Genetics:
A;Genetics:
A;Genetics:
A;Genetics:
A;Genetics:
A;Genetics:
A;Genetics:
A;Genetics:
A;Genetics:
A;Genetics:
A;Genetics:
A;Genetics:
A;Genetics:
A;Genetics:
A;Genetics:
A;Genetics:
A;Genetics:
A;Genetics:
A;Genetics:
A;Genetics:
A;Genetics:
A;Genetics:
A;Genetics:
A;Genetics:
A;Genetics:
A;Genetics:
A;Genetics:
A;Genetics:
A;Genetics:
A;Genetics:
A;Genetics:
A;Genetics:
A;Genetics:
A;Genetics:
A;Genetics:
A;Genetics:
A;Genetics:
A;Genetics:
A;Genetics:
A;Genetics:
A;Genetics:
A;Genetics:
A;Genetics:
A;Genetics:
A;Geneti
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     protein YKR029c homolog YJL105w - yeast (Saccharomyces cerevisiae)
N:Alternate names: hypothetical protein J0819
C;Speciaes: Saccharomyces cerevisiae
C;Date: 05-May-1995 #sequenc_revision 01-Sep-1995 #text_change 21-Jul-2000
C;Accession: 553382; 556883; 557363
R;Rasmussen, S.W.
Submitted to the EMBL Data Library, Pebruary 1995
A;Description: A 37.5 kb region of yeast chromosome X includes the SME1, MEF2, GSH1 and A;Reference number: 553376
A;Accession: 553382
A;Accession: 553382
A;Accession: 553382
A;Goos references: EMBL:X85021; NID:g728698; PID:g728705
R;Rasmussen, S.W.
submitted to the Protein Sequence Database, September 1995
A;Reference number: 556876
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  421 KYNKEALSPPRRKAFKKWTPPRSPFNLVQETLFHDPWKLLIATIFLNRTSGKMAIPVLWK 480
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   69 KSVKKQKKSKK---KEESSSESESESESSSSESESESSSSS---ESESSSSESSSES 122
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              305 EENSLVK---KKERSLSSGSNFCSEQKTSGIINKFCSAKDSEHNEKYEDTFLESEEIGTK 361
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             123 EEEVIVKTEEKKESSSESSSSESEEEEEAVV-KIEEKKESSSDSSSESSSESESESSS 181
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        362 VEVVERKEHLH-TDILKRGSEMDNNCSPTRKDFTGEKIFQEDTIPRTQIERRKTSLYFSS 420
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   182 SESEEEEEVVEKTEEKKEGSSESSDDSESSDDSSSE---SGDSDSSSDSESESS---SE 234
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            235 DEKKRKAEPASEERPAKITKPSQDSN---ET------CTVFVGRLSWNVDDQWLGQ 281
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   185 KKDVFMPPSSSSELQESRGLSNFTSTHLLLKEDEGVDDVNFRKVRKPKGKVTILKGIPIK 244
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        245 KTKKGCRKSCSGFVQSDSKRESVCNKADAESEPVAQKSQLDRTVCISDAGACGETLSVTS 304
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            Gaps
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         19 KKGAIEKPSKSKKITKEAAKE-----IAKQSSKTDVSPKKSKKEAKRASSPE--PSK 68
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        41;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 4.5%; Score 138.5; DB 2; Length 500; 23.7%; Pred. No. 0.16; tive 50; Mismatches 147; Indels 41
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        74; Conservative
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    481 FLEKYPSAEVAR 492
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          282 EFEEYGTIVGAR 293
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              Best Local Similarity
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             Query Match
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 •
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            Matches
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  q
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   q
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         qq
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   qq
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      ŏ
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               ŏ
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              qq
```

```
qq
                                                                                                                                                                                                                                                                                                                                              a
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    qq
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                qq
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          Q
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             ó
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           qa
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  Qγ
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                q
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   ŏ
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        qq
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            qq
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               δò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             QQ
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    δŏ
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      q
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         δò
                                                                                           ŏγ
                                                                                                                                                                                                                                                                δ
                                                                                                                                                                                                                                                                                                                                                                                                                                 δy
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            ŏ
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       A; Residues: 1-1359 <FUL>
A; Cross-references: EMBL:AF000196; PIDN:AAC24256.1; GSPDB:GN00019; CESP:B0041.7
A; Experimental source: strain Bristol N2; clone B0041
C; Genetics:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             19;
                                                                                                                                                                                                                                                                21;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     C;Species: Caenorhabditis elegans
C;Date: 29-Oct-1999 #sequence_revision 29-Oct-1999 #text_change 29-Oct-1999
C;Accession: T34036
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         172 KRSKTSNDPISEICSENSL---NDFLMLVGQ-----SDDESINFDAMLRNYSQNNSKN 222
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  223 YSLKYTKYTKKDTKKNAKKSSDAESEDSKDAKKDSKKVKKNVKKDDKKKDVKKDTESTDA 282
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               --GCRKSCSGFVQS 260
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                DSKRES--VCNKADAESEPVAQKSQLDRTVCISDAGACGETLSVTSEENSLVKKKERSLS 318
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 SGSNFCSEQKTSGIINKFCSAKDSEHNEKYEDT -- FLESEEIGTKVEVVERKEHLHTDIL 376
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 393 ADSEPKGDSK-KGKKDEKKGKKDSKKDDKKKDAKKNAESTEMESDLELKKDKKHSKE--- 448
                                                                                                                                                                                                                                                                                                                                                                                                                             70 GGTPLKKDSKKKGGSYATNPESKQIVEEKTKRQNEAD-KTPLKSSHENEQSKKSKSKSSET 128
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    62 NPLLQEPIASAQFGATAGTECRKSVPCGWERVVKQRLFGKTAGRFDVYFI----SPQGLK 117
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                118 FRSKSSLANYLHKNGETSLKPEDFDFTVLSKRGIKSRYKDCSMAALTSHLQN--QSNNSN 175
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          WNLRTRSKCKKDV -- FMPPSSSSELQESR-GLSNFTSTHLLLKEDEGVDDV ---- 223
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         283 ESGDSKDERKDTK-KDKKKLKKDDKKKDTKKYPESTDTESGDAKDARNDSRNLKKASKND 341
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          342 DKKKDAKKITFSTDSESELESKESQKDEKKDKKD-----SKTDNKKS-VKNDEESTD 392
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                75 GATAGTECRKSVPCGWERVVKQRLFGKTAGRFDVYFISPQGLKFRSKSSLANYLHKNGET 134
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              NP-----ESQNSKTVSKNCSQ------KDKKDSKNSKKTNTEFLHTKNNPKKDL 171
                                                                                                                                                                                                                                                                    Gaps
                                                                                                                                                                                                                                                                                                                                              2 GTTGLESLSLGDRGAAPTVTSSERLVPDPPNDLRKEDVAMELERVGEDEEQMMIKRSSEC 61
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             91; Mismatches 253; Indels 131; Gaps
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           15 GAAPTVTSSERLVPDPPNDLRKEDVAMELERVGEDEEQMMIKRSSECNPLLQEPIASAQF 74
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        | CONTROL | CONTROL | CONTROL | CONTROL | CONTROL | CONTROL | CONTROL | CONTROL | CONTROL | CONTROL | CONTROL | CONTROL | CONTROL | CONTROL | CONTROL | CONTROL | CONTROL | CONTROL | CONTROL | CONTROL | CONTROL | CONTROL | CONTROL | CONTROL | CONTROL | CONTROL | CONTROL | CONTROL | CONTROL | CONTROL | CONTROL | CONTROL | CONTROL | CONTROL | CONTROL | CONTROL | CONTROL | CONTROL | CONTROL | CONTROL | CONTROL | CONTROL | CONTROL | CONTROL | CONTROL | CONTROL | CONTROL | CONTROL | CONTROL | CONTROL | CONTROL | CONTROL | CONTROL | CONTROL | CONTROL | CONTROL | CONTROL | CONTROL | CONTROL | CONTROL | CONTROL | CONTROL | CONTROL | CONTROL | CONTROL | CONTROL | CONTROL | CONTROL | CONTROL | CONTROL | CONTROL | CONTROL | CONTROL | CONTROL | CONTROL | CONTROL | CONTROL | CONTROL | CONTROL | CONTROL | CONTROL | CONTROL | CONTROL | CONTROL | CONTROL | CONTROL | CONTROL | CONTROL | CONTROL | CONTROL | CONTROL | CONTROL | CONTROL | CONTROL | CONTROL | CONTROL | CONTROL | CONTROL | CONTROL | CONTROL | CONTROL | CONTROL | CONTROL | CONTROL | CONTROL | CONTROL | CONTROL | CONTROL | CONTROL | CONTROL | CONTROL | CONTROL | CONTROL | CONTROL | CONTROL | CONTROL | CONTROL | CONTROL | CONTROL | CONTROL | CONTROL | CONTROL | CONTROL | CONTROL | CONTROL | CONTROL | CONTROL | CONTROL | CONTROL | CONTROL | CONTROL | CONTROL | CONTROL | CONTROL | CONTROL | CONTROL | CONTROL | CONTROL | CONTROL | CONTROL | CONTROL | CONTROL | CONTROL | CONTROL | CONTROL | CONTROL | CONTROL | CONTROL | CONTROL | CONTROL | CONTROL | CONTROL | CONTROL | CONTROL | CONTROL | CONTROL | CONTROL | CONTROL | CONTROL | CONTROL | CONTROL | CONTROL | CONTROL | CONTROL | CONTROL | CONTROL | CONTROL | CONTROL | CONTROL | CONTROL | CONTROL | CONTROL | CONTROL | CONTROL | CONTROL | CONTROL | CONTROL | CONTROL | CONTROL | CONTROL | CONTROL | CONTROL | CONTROL | CONTROL | CONTROL | CONTROL | CONTROL | CONTROL | CONTROL | CONTROL | CONTROL | CONTROL | CONTROL | CONTROL | CONTROL | CONTROL | CONTROL | CONTROL | CONTROL | CONTROL | CONTROL | CONTROL | CONTROL | CONTROL | CONT
A,Cross-references: GB:222780; NID:9396104; PIDN:CAA80457.1; PID:9396105 C;Keywords: cytoskeleton
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 'Map position: 1
'Introns: 12/2; 59/2; 248/2; 582/1; 675/1; 733/3; 966/2; 1044/2; 1310/3
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      377 KRGSEMDNNCSPTRKDFTGEKIFQEDTIPRTQIERRKTSLYFSSKYNKEALSPPRRK 433
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 Query Match
4.5%; Score 136.5; DB 2; Length 1359;
Best Local Similarity 19.2%; Pred. No. 0.82;
Matches 113; Conservative 91; Mismatches 253; Indels 131;
                                                                                                                                                                                                                                                                Indels
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         R;Fulton, R.; Wohldmann, P. submitted to the EMBL Data Library, April 1997
A;Description: The sequence of C. elegans cosmid B0041.
A;Reference number: Z21466
A;Recession: T34036
                                                                                                                                                                         Query Match 4.5%; Score 136.5; DB 2; Best Local Similarity 23.3%; Pred. No. 0.27; Matches 111; Conservative 72; Mismatches 205;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             hypothetical protein B0041.7 - Caenorhabditis elegans
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      A;Status: preliminary; translated from GB/EMBL/DDBJ A;Molecule type: DNA
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               -----NFRKVRKPKGKVTILKGIPIKKTKK-
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              Gene: CESP:B0041.7
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  129
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          176
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           224
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              셤
                                                                                                                                                                                                                                                                                                                                              ò
                                                                                                                                                                                                                                                                                                                                                                                                                             g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                à
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               Dp
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       οy
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      Q
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 à
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            q
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    á
```

```
A; Cross-references: EMBL: AL122011; PIDN: CAB58409.1; GSPDB: GN00068; SPDB: SPCC825.01
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         14;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            probable ABC transporter - fission yeast (Schizosaccharomyces pombe)
C;Species: Schizosaccharomyces pombe
C;Date: 03-Dec-1999 #sequence_revision 03-Dec-1999 #text_change 03-Dec-1999
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          C;Accession: T41622
R;Aert, R; Volckaert, G.; McDougall, R.C.; Rajandream, M.A.; Barrell, B.G.
submitted to the EMBL Data Library, October 1999
A;Reference number: 221735
A;Scession: T41622
A;Status: preliminary; translated from GB/EMBL/DDBJ
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     -----VEVLPQKK-----KRGAVTLISDSEDEKDQKSESEASDVEEKVSKKKAK 342
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      396 KERR-----KRLEKKQKEFNGIVLEEGEDLTEMLTGTSSQRKLKSVVLDPDSSTVDEESK 450
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  266 SVCNKADAE-----SEPVAQKSQLDRTVC1SiAGACGETLSVTSEENSLVKKKERS 316
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  317 LSSGSNFCSEQKTSGIINKFCSAKDSEHNEKYEDTFLESEEIGTKVEVVERKEH----LH 372
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     302 VTSEENSLVKKKERSLSSGSNFCSEQKTSG-----IINKFCSAKDSEHNEKYEDTF 352
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           353 LESEEIGTKVEVVERKEHLHTDILKRGSEMDNNCSPTRKDFTGEKIFQ--EDTIPRTQIE 410
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      411 RRKTSLYFSSKYNKEALSPPRRKAFKKWTPPRSPFNLVQETLFHDPWKLLIATIFLNRTS 470
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               343 KQESS--ESGSDSSEGSITVNRKSKKKEKPEKKKKGIIM-----DSSKLQKETIDAERAE 395
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  149 RGIKS---RYKDCSMAALTSHLQNQSNNSNWNIRTRSKCKKDVFMPPSSSSELQESRGLSN 206
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         207 FTSTHLLLK-EDEGVDDVNFRKVRKPKGKVTLLKGIPIKKTKKGCRKSCSGFVQSDSKRE 265
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 101 E--EEEDSEKPVRKNKKSSKKASPRNAFDALADDMDDLSLDEEESESSEKSKKKKKSKS 158
                                                                             135 SLKPEDFDFTVLSKRGIKSRYKDCSMAALTSHLQNQSNNSNWNLRTRSKCKKDVFMPPSS 194
                                                                                                                                                  95 SDEEEDRKKSKSKKKVDQKKKEKSKKKRTTSSSEDEDSDEEREQKSKKKSKKTKKQTSSE 154
                                                                                                                                                                                                                                                                                                             ------RKVKKSKKNKEKSVKKRAETSEESDEDEK 191
------REGKPPPKKRPAKKRKASSSEEDDDDEEESPR--KSSKKSRKRAKSESESDE 94
                                                                                                                                                                                                                                                                                                                                                                                          242 PIKKTKKGCRKSCSGFVQSDSKRESVCNKADAESEPVAQKSQLDRTVCISDAGACGETLS
                                                                                                                                                                                                                                                                                                                                                                                                                                                                       192 PSKKSKKGLKKKAKSESSESEDEKEVKKSKKKSKKVVKKES------ESEDEAPEKKK
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            3 RGKRSTQRDADLELESLQSETESESPQP----VTKSKAKK-----NKKKLNKASAFNS
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          245 TEKRKRSKTSSEESSESSEKSDEEEERKESSPKPKKKKPLAVKKLSSDEESEESD-----
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             471 GKMAIPVLWKFLE-KYPSAEVARTADWRDVSELLKPLGLYDLRAKTIVKFSDEYLTKQWK
                                                                                                                                                                                                                               SSELQESRGLSNFTSTHLLLKEDEGVDDVNFRKVRK-PKGKVTILK------GI
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         :06
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        4.4%; Score 135; DB 2; Length 822;
22.4%; Pred. No. 0.52;
tve 53; Mismatches 134; Indels
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        530 YPIELHGIGKYGNDSYRIFCVNEWKQVHPEDHKLNKYHDWLWENHEKL 577
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    A; Experimental source: strain 972h-; cosmid c825
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   ..
-::
-:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        Query Match 4.4%;
Best Local Similarity 22.4%;
Matches 80; Conservative 5
                                                                                                                                                                                                                                                                                                             155 SSEESEEE-----
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      1-822 <AER>
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 A;Gene: SPDB:SPCC825.01
A;Map position: 3
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           A; Residues:
                                                                                                                                                                                                                                   195
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         299
```

Page

```
A; Cross-references: EMBL: AL121859; PIDN: CAB58376.1; GSPDB: GN00068; SPDB: SPCP31B10.10c
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         A;Cross-references: EMBL:AL031323; PIDN:CAA20434.1; GSPDB:GN00068; SPDB:SPCC962.02c A;Experimental source: strain 9724-; cosmid c962 K:Wedler, H.: Duesterhoeft. A.: McDougall, R.C.: Rajandream, M.A.; Barrell, B.G. submitted to the EMBL pata Library, October 1999
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              cuti7 protein - fission yeast (Schizosaccharomyces pombe)
C;Species: Schizosaccharomyces pombe
C;Species: Schizosaccharomyces pombe
C;Species: 21-Jan-2000 #sequence_revision 21-Jan-2000 #text_change 02-Jun-2000
C;Accession: T43523; T41649; T41700
R;Morishita, J.; Matsusaka, T.: Yanagida, M.
submitted to the EMBL Data Library, August 1999
A;Description: Fission yeast cuti7 is required for chromosome segregation.
A;Relerence number: 222536
                                                                                                                                         1062 LLTPEARKLLEEAKESVKAYKDCVSRAR-NEKEKQECEKLLTPEARKLLEQEVKKSVKAY 1060
                                                                                                                                                                                                                                                                               1061 LDCVSRARNEKEKQECEKLLTPEARKLLENQALD--CLKNAKTEAEKKRCVKDLPKDLQK 1118
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  1162 --KSVKAYLDCVSKAKNEAEKKECEKLLTPEARKLLEEAKESVKAYKDCVSRARNEKEKQ 1219
   942 LENQALDCLKNAKTEAEKKRCVKDLPKDLQKKVLAKESVRVYLDCVSKAKNEAERKECEK 1001
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       1220 ECEKLLTFEARKLLEQEVKKSVKAYLDCVSRARNEKEKQECEKLLTPEARKFLEKQRQQK 1279
                                                                                                                                                                                                                                                                                                                                                                                                                            1119 KVLAKESVKA-YLDC----VSRARNE-----KEKKECEK--LLTPEARKLLEESK- 1161
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 1280 DKAIKDCLKNADPNDRAAIMKCLDGLSDEEKL,------KYLQEAREKAVLDCLK 1327
                                                                      50 ----EEQMMIKRSSECNPLLQEPIASAQFGATAGTECRKSVPCGWERVVKQRLFGKTAGR 105
                                                                                                                                                                                                                -----FISPOGLKFRSKSSLANYLHKNGETSLK-----PEDFDF 143
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           Query Match 4.3%; Score 131.5; DB 2; Length 997;
Best Local Similarity 20.1%; Pred. No. 1.1;
Matches 129; Conservative 119; Mismatches 241; Indels 153; Gaps
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        -----RKSCSGFVQSDSKRESVCNKADAES--EPVA-----QKSQL 284
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              285 DRTV--CI-----SDAGACGETLSVISEENSLVKKKERSLSSGSNFCSEQKTSGIINKFC 337
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        338 SAKDSEHNEKYEDTFLE-SEEIGTKVEVVERKEHLHTDILKRGSE-MDNNCSPTRKD 392
                                                                                                                                                                                                                                                                                                                                                     144 TVLSKRGIKSRYKDCSMAALTSHLQNQSNNSNWNLRTRSKCKKDVFMPPSSSSELQESRG
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               204 LSNFTSTHL--LLKEDEGVDDVNFRKVRKPKGKVTILKGIPIKKTKKGC--
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       A;Molecule type: mRNA
A;Residues: 1-997 <MOR>
A;Cross-references: EMBL:AB031034; PIDN:BAA83415.1
A;Cross-references: EMBL:AB031034; PIDN:BAA83415.1
R;Harris, D; Wood, V; Rajandream, M.A.; Barrell, B.G. submitted to the EMBL Data Library, August 1998
A;Reference number: Z22007
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                A; Gene: cut17; SPCC962.02c; SPDB:SPCP31B10.10c
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 A;Status: translated from GB/EMBL/DDBJ
A;Molecule type: DNA
A;Residues: 1-997 <HAR>
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      A;Status: translated from GB/EMBL/DDBJ
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         A; Status: translated from GB/EMBL/DDBJ
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          A; Reference number: 222010
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                <WED>
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           A; Molecule type: DNA A; Residues: 932-997 <
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       A; Accession: T41649
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        A; Accession: T41700
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        A; Accession: T43523
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             A; Map position: 3L
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  RESULT 13
                                                                  δ
                                                                                                                                         qq
                                                                                                                                                                                                                δă
                                                                                                                                                                                                                                                                               g
                                                                                                                                                                                                                                                                                                                                                     ŏ
                                                                                                                                                                                                                                                                                                                                                                                                                        Op
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              q
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                qq
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              ŏλ
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 q
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           δŏ
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      q
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           ŏ
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    \tilde{Q}
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        A; Variety: strain J99
C; Date: 12-Feb-1999 #sequence_revision 12-Feb-1999 #text_change 08-Oct-1999
C; Date: 12-Feb-1999 #sequence_revision 12-Feb-1999 #text_change 08-Oct-1999
C; Accession: A71928
R; Alm, R.A.; Ling, L.S.L.; Moir, D.T.; King, B.L.; Brown, E.D.; Doig, P.C.; Smith, D.R.; Ives, C.; Gibson, R.; Merberg, D.; Mills, S.D.; Jiang, Q.; Taylor, D.E.; Vovis, G.F.; Nature 397, 176-180, 1999
A; Title: Genomic sequence comparison of two unrelated isolates of the human gastric path A; Reference number: A71800; MUID:99120557
A; Accession: A71928
A; Accession: A71928
A; Molecule type: DNA
A; Residues: 1-1819 < ARN>
A; Cross references: GB: AE001481; GB: AE001439; NID:94155005; PIDN: AAD06047.1; PID:9415500
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    R:Kawarabayasi, Y.; Hino, Y.; Horikawa, H.; Yamazaki, S.; Haikawa, Y.; Jin-no, K.; Takahawa, H.; Takamiya, M.; Masuda, S.; Funahashi, T.; Tanaka, T.; Kudoh, Y.; Yamazaki, J.; Kanaha, R.; S. 6, 83-101, 1999
A;Title: Complete genome sequence of an aerobic hyper-thermophilic Grenarchaeon, Aeropyr A;Reference number: A72450; MUID:99310339
A;Accession: A72682
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       A;Cross-references: DDBJ:AP000060; NID:95104188; PIDN:BAA79857.1; PID:d1043643; PID:9510
                                                                                                                                                                                                                                                                                                                                                                                                                                                      probable A/G-specific adenine glycosylase APE0875 - Aeropyrum pernix (strain K1)
C;Species: Aeropyrum pernix
C;Date: 20-Aug-1999 #sequence_revision 20-Aug-1999 #text_change 20-Aug-1999
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          ω,
159 KDDG----SEALDDGDI-----ESSEKDKKKKKKKKRENDDAPKKDRKTRKKEEKARKLA 208
                                                                      -----FTGEKIFQEDTIPRTQIERRKTSL 416
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  209 SMLESENKDNDANAAPLNKTDAFKDGLPSGRLIFAYASGQKVAPDGSNPADGI----- 261
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      455 DPWKLLIATIFLNRTSGKMAIPVLWKFLEKYPSAEVARTADWRDVSELLKPLGLYDLRAK 514
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          Gaps
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   Gaps
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          6 LESLSLGDRGAAPTVTSSERLVPDPPNDLRKEDVAMELERVGED-------49
                                                                                                                                                                                                        417 YFSSKYNKEALSPPRRKAFKKWTPPRSPFNLVQETLFHDPW-KLLIATIFLNRTSGK 472
                                                                                                                                                                                                                                                                               262 ----TVTGNLLSPPNSR------DLQVEKLSVSAWGKLLIKDSELNLINGR 302
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  18.2%; Pred. No. 2.4;
Live 83; Mismatches 183; Indels 124;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            4.3%; Score 132.5; DB 2; Length 286; 29.8%; Pred. No. 0.18; tive 26; Mismatches 31; Indels 9.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        4.3%; Score 132; DB 2; Length 1819;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         cag island protein - Helicobacter pylori (strain J99)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               515 TIVKFSDEYLTKQW--KYPI-----ELHGIGKY 540
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   158 HLIELA-KHIEARYGGRIPCSKEKLKELPGVGDY 190
                                                                  373 TDILKRGSEMDNNCSPTRKD-----
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             A; Experimental source: strain J99
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     A; Experimental source: strain K1 C; Genetics:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 Best_Local Similarity 29.89
Matches 28; Conservative
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               Best Local Similarity 18.2%
Matches 87; Conservative
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           Species: Helicobacter pylori
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             A; Molecule type: DNA
A; Residues: 1-286 <KAW>
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               preliminary
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              C; Accession: A72682
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             A; Gene: orf13/14
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            A; Gene: APE0875
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    Query Match
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        Query Match
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          RESULT 12
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               Genetics
                                                                                                                                  qq
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          ŏλ
                                                                  δ
                                                                                                                                                                                                        δλ
                                                                                                                                                                                                                                                                       Dp
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         á
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         qq
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               δ
```

20 VTSSERIVPDPPNDLRKEDVAMELERVGEDEEOMMIKRSSECNPLLQEPIASA 72	QY 68 PIASAQFGATAGTECRKSVPCGWERVVKQRLFGKTAGRFDVY 109
	Db 1194 LLEBARESLKAYKDCLSQARNEEERRACEKLLTPEARKLLEQEVKKSIKAYLDCVSRARN 1253
y 73 QFGATAGTECRKSVPCGWERVVKQRLFGKTAGRFDVYFISPQGLKFRSKSSLANYLHKNG 132	QY 110FISPQGLKFRSKSSLANYLHKNGETSLKPEDFDFTVLSKRGIKS 153 :: : : : :
133 ETSLKPEDFDFTVLSKRGIKSRYKDC 	QY 154 RYKDCSMAALTSHLQNQSNNSNWNLRTRSKCKKDVFWPPSSSSELQESRGLSNFTSTHLL 213
y 174 SNWNLRTRSKCKKDVFMPPSSSELQESRGLSNFTSTHILLKEDEGVDD 222	QY 214 LKEDEGVDDVNFRKVKRP-KGKVTILKGIPIKKTKKGCRKSCSGFVQSDSKRESVCNKAD 272 : : : : : : Dayler - - - - - - - - - - - -
223 VNFRKVRRPKGKVTILKGIPIKKTKKGCRKSCSGEVQSDSKRESV 267 223 VNFRKVRRPKGKVTILKGIPIKKTKKGCRKSCSGEVQSDSKRESV 267 224 FIEKPKTEISEVLPPERKRKAICDESQTVRVSIDRGVTKTRDVSSPVSDEKSENV 697	QY 273 AESEPVAQKSQLDRTVCISDAGACGETLSVTSEENSLVKKERSLSS 319
2y 268 CNKADAESEPVAQKSQLDRTVCISDAGACGETLSVTSEENSLVKK 312 : : : : : : : :	OY 320 GSNFCSEOKTSGIINKFCSAKDSEHNEKYEDTFLE-SEEIGTKVEVVERKEHLHTDILKR 378 1
313 KERSLSSGSNFCSEQKTSGIINKFCSAKDSEHNEKYEDTFLESEEIGTKVEVVERKEHLH 372 1	Oy 379 GSE-MDNNCSPTRKD 392 : ::
2y 373 TDILKRGSEMDNNCSPTROFTGEKIFQEDTIPRTQIERRKTSLYFSSKY-NKEAL 427 :: :: :	RESULT 15 A96748 Notestation of the state of
428 SPPRKKAFKKWTPPRSPFNLVQETLFHDPWKLLIATIFLNRTSGKMAIPVLWKFL 482	hypotnetical protein indiger. I improved a marginal constraint () Species. Arabidopsis thaliana (mouse-ear cress) () Species. Arabidopsis thaliana (mouse-ear cress) () Species. 02-Mar-2001 #sequence_revision 02-Mar-2001 #text_change 31-Mar-2001 () Accession: A96748
483 EKYPSAEVARTADWRDVSELLKPLGLYDLRAKTIVKFSDEYLTKOW-KYPIEL 534 9-00	Rifneologis, A.; Ecker, J.K.; Palm, C.J.; Federspiel, N.A.; Naul, S.; Miles, C.; Palm, C.W.; Chung, M.K.; Conn, L.; Conway, A.B.; Conway, A.R.; Creasy, T.H.; Dewar, ansen, N.F.; Hughes, B.; Hulzar, L. Nature 408, 816-820, 2000
535 HGIGKYGNDSYRIEVNEWCOVPPDHKLMKYHDWLWENHEK 576 11	A; Authors: Hunter, J.L.; Jenkins, J.; Johnson Hopson, C.; Khan, S.; Khaykin, E.; Kim, C.A.; Li, J.H.; Li, Y.; Lin, X.; Liu, S.X.; Liu, Z.A.; Luros, J.S.; Maiti, R.; Marzia Rizzo, M.; Rooney, T.; Rowley, D.; Sakano, H. A; Authors: Salzaberg, S.L.; Schwartz, J.R.; Shinn, P.; Southwick, A.M.; Sun, H.; Tallo A; Authors: Salzaberg, S.L.; Schwartz, J.R.; Shinn, P.; Southwick, A.M.; Sun, H.; Tallo Liv M. G. Praser, C.M.; W. Norther, J.C.; Davis, R.W.
RESULT 14	A; Title: Sequence and analysis of chromosome 1 of the plant Arabidopsis. A; Reference number: A86141; MUID:21016719 A, Reference number: A86141; MUID:21016719
thogenicity island protein cag7 - Helicobacter pylori (strain 26695) les: Helicobacter pylori . 09-Aug-1997 #sequence_revision 09-Aug-1997 #text_change 08-Oct-1999	A. Status: preliminary A. Molecule type: DNA A. Rolecule type: DNA A. Residues: 1-946 < STO A. CREATOR A. ROLECULE TO STORY A. CREATOR A. ROLECULE TO STORY A. CREATOR A. CREATO
	C;Genetics: C;Genetics: A;Gene: T10D10.13 A;Map position: 1
Asture 388, 339-347, 1397. A; Authors: Wallin, E.; Hayes, W.S.; Borodovsky, M.; Karpk, P.D.; Smith, H.O.; Fraser, C. A; Title: The complete genome sequence of the gastric pathogen Helicobacter pylori. A; Reference number: A64520; MUID:97394467	Query Match 4.2%; Score 128.5; DB 2; Length 946; Best Local Similarity 21.0%; Pred. No. 1.7; Matches 94; Conservative 56; Mismatches 177; Indels 121; Gaps 16;
A;Status: preliminary; nucleic acid sequence not shown; translation not shown A;Molecule type: DNA A;Molecule type	EDVAMELERVGEDEEQN : : LFAEHQLRVPGDQSSSE
Query Match 4.3%; Score 130; DB 2; Length 1927; Best Local Similarity 20.5%; Pred. No. 3.5;	QY 90 -WERVUKQRLFGKTAGRFDVYFISPQGLKFRSKSSLANYLHKNGETSLKPEDFDFTVLSK 148 :::::: ::
Matches 89; Conservative 61; Mismatches 147; Indels 138; Gaps 21; Qy 17 APTVTSSERLVPDPPNDLRKEDVAMELERVGEDEEQMMIKRSSECNPLLQE 67	OY 149 RGIKSRYKDCSMAALTSHLQNQSNNSNWNLRTRSKCKKDVFMPPSSSSELQESRGLSNFT 208 : 1 : 1 : 1 : 1 : 1 1 : 1 1 : 1 1
Db 1137 AKTEADKKRCVKDLPKDLQKKVLAKESVKAYLDCVSRARNEKEKKECEKLLTPEAKK 1193	209

: : : 531 SRSSMKKYQSEBDEDISEQKPRAKDKAASGQQSVGSISSRSSQARKLQPNRNMSSSITPR 590	LKGIPIKKŢKKĢCŖKŚCSGFVÇ	SAASVPKPSGKVSNTSSGRRRSDKSLAQSVPNFSELIKENTKPSSL 636	284 LDRTVCISDAGACGETLSVTSEENSLVKKKERSLSSGSNFCSEQKTSGIIN 334	637 AVKTTMRSQVKSSGRTKNIKEDTLLQRPRSLRKSSSGNIDFTELSTLCSDDMMVS 691	335 KFCSAKDSEHNEKYEDTFLESEEI-GTKVEVVERKEHLHTDILKRGSEM 382	692 LRVDSDISETLRNEEYDEPEAEPEEVLENAVREEEEVEELETLVFEDGNPMLSEAYEKVD 751	PTRKDFTGEKI	752 HSGEBNCSFLPATVPTT 768	
531	227	591	284	1 1 1 1	335 1	692 1	383 -	752 1	
qa	QY	qq	Qy	Ф	QY	DÞ	Qy	QQ	

Search completed: January 15, 2002, 13:52:37 Job time: 406 sec

```
Sequence 2, Appli
Sequence 3, Appli
Sequence 3, Appli
Sequence 11, Appl
Sequence 2, Appli
Sequence 6, Appli
Sequence 6, Appli
Sequence 6, Appli
Sequence 12, Appli
Sequence 12, Appl
Sequence 12, Appli
Sequence 12, Appli
Sequence 34, Appli
Sequence 34, Appli
Sequence 34, Appli
Sequence 34, Appli
Sequence 6, Appli
Sequence 18, Appli
                                                                                                                     Search time 57.72 Seconds (without alignments) 123.589 Million cell updates/sec
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             Sequence 2, Appli
Sequence 12, Appl
Patent No. 5349058
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.
                                                                                                                                                                                                      US-09-724-296-37
1662
1 MIFRFGFVSNAMSLWDASPA......ELSSIRGVKRIGGGALQWKS 317
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             Description
                                                                                                                                                                                                                                                                                                                                                                                                  212252
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        /cgg2_6/ptodata/2/iaa/5A_COMB.pep:*
/cgn2_6/ptodata/2/iaa/5B_COMB.pep:*
/cgn2_6/ptodata/2/iaa/6A_COMB.pep:*
/cgn2_6/ptodata/2/iaa/6B_COMB.pep:*
/cgn2_6/ptodata/2/iaa/PCTUS_COMB.pep:*
/cgn2_6/ptodata/2/iaa/PCTUS_COMB.pep:*
GenCore version 4.5 Copyright (c) 1993 - 2000 Compugen Ltd.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              hits satisfying chosen parameters:
                                                                                                                                                                                                                                                                                                                                                        212252 seqs, 22503292 residues
                                                                                                                        January 15, 2002, 13:51:13;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              SUMMARIES
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries
                                                                                OM protein - protein search, using sw model
                                                                                                                                                                                                                                                                                           BLOSUM62
Gapop 10.0 , Gapext 0.5
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           Issued_Patents_AA:*
                                                                                                                                                                                                                                                                                                                                                                                                                                       Minimum DB seq length: 0
Maximum DB seq length: 2000000000
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             DB
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  Query
Match Length
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        Post-processing:
                                                                                                                                                                                                                                                                                                                                                                                                  Total number of
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  993.5

982.5

982.5

983.4

983.5

983.5

983.5

983.5

983.5

983.5

983.5

983.5

983.5

983.5

983.5

983.5

983.5

983.5

983.5

983.5

983.5

983.5

983.5

983.5

983.5

983.5

983.5

983.5

983.5

983.5

983.5

983.5

983.5

983.5

983.5

983.5

983.5

983.5

983.5

983.5

983.5

983.5

983.5

983.5

983.5

983.5

983.5

983.5

983.5

983.5

983.5

983.5

983.5

983.5

983.5

983.5

983.5

983.5

983.5

983.5

983.5

983.5

983.5

983.5

983.5

983.5

983.5

983.5

983.5

983.5

983.5

983.5

983.5

983.5

983.5

983.5

983.5

983.5

983.5

983.5

983.5

983.5

983.5

983.5

983.5

983.5

983.5

983.5

983.5

983.5

983.5

983.5

983.5

983.5

983.5

983.5

983.5

983.5

983.5

983.5

983.5

983.5

983.5

983.5

983.5

983.5

983.5

983.5

983.5

983.5

983.5

983.5

983.5

983.5

983.5

983.5

983.5

983.5

983.5

983.5

983.5

983.5

983.5

983.5

983.5

983.5

983.5

983.5

983.5

983.5

983.5

983.5

983.5

983.5

983.5

983.5

983.5

983.5

983.5

983.5

983.5

983.5

983.5

983.5

983.5

983.5

983.5

983.5

983.5

983.5

983.5

983.5

983.5

983.5

983.5

983.5

983.5

983.5

983.5

983.5

983.5

983.5

983.5

983.5

983.5

983.5

983.5

983.5

983.5

983.5

983.5

983.5

983.5

983.5

983.5

983.5

983.5

983.5

983.5

983.5

983.5

983.5

983.5

983.5

983.5

983.5

983.5

983.5

983.5

983.5

983.5

983.5

983.5

983.5

983.5

983.5

983.5

983.5

983.5

983.5

983.5

983.5

983.5

983.5

983.5

983.5

983.5

983.5

983.5

983.5

983.5

983.5

983.5

983.5

983.5

983.5

983.5

983.5

983.5

983.5

983.5

983.5

983.5

983.5

983.5

983.5

983.5

983.5

983.5

983.5

983.5

983.5

983.5

983.5

983.5

983.5

983.5

983.5

983.5

983.5

983.5

983.5

983.5

983.5

983.5

983.5

983.5

983.5

983.5

983.5

983.5

983.5

983.5

983.5

983.5

983.5

983.5

983.5

983.5

983.5

983.5

983.5

983.5

983.5

983.5

983.5

983.5

983.5

983.5

983.5

983.5

983.5

983.5

983.5

983.5

983.5

983.5

983.5

983.5

983.5

983.5

983.5

983.5

983.5

983.5

983.5

983.5

983.5

983.5

983.5

983.5

983.5

98
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             Score
                                                                                                                                                                                                                              Perfect score:
                                                                                                                                                                                                                                                                                           Scoring table:
                                                                                                                                                                                                                                                      Sequence:
                                                                                                                                                                                                                                                                                                                                                        Searched:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           Database
                                                                                                                          Run on:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   Result
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             .
9
```

28 75 4.5 704 1 US-08-188-582-5 30 75 4.5 704 1 US-08-188-582-5 31 74.5 4.5 1257 3 US-08-09-17-63-5 31 74.5 4.5 369 1 US-08-09-17-63-5 31 74.5 4.5 369 1 US-08-09-17-63-5 31 74.5 4.5 381 1 US-08-09-17-83-5 34 74.5 4.5 381 1 US-08-09-18-2 35 74.5 4.5 381 1 US-08-09-18-2 36 74.5 4.5 381 1 US-08-09-18-2 37 74.5 4.5 381 1 US-08-09-18-18-2 38 74.5 4.5 976 4 US-09-11-9-16-2 40 74 4.5 976 4 US-09-11-9-16-2 41 74 4.5 976 4 US-09-19-10-11 42 74 4.5 976 4 US-09-19-10-11 42 74 4.5 976 4 US-09-19-10-11 43 74 5 4.5 263 4 US-09-19-10-11 44 7 4 4.5 976 4 US-09-19-10-11 45 73.5 4.4 264 2 US-08-19-10-11 48 74 4.5 2958 4 US-08-19-10-11 49 74 4.5 104-13-18-2 45 73.5 4.4 264 2 US-08-10-213-2 45 73.5 4.4 2 264 2 US-08-10-213-2 45 73.5 4.4 2 264 2 US-08-10-213-2 45 73.5 4.4 2 264 2 US-08-10-213-2 45 73.5 4.4 2 264 2 US-08-10-2 45 73.5 4.4 2 64 2 US-08-10-2 45 73.5 4.4 2 64 1 US-08-10-2 45 73.5 4.4 2 64 1 US-08-10-2 45 73.5 4.4 2 64 1 US-08-10-2 45 73.5 4.4 2 64 1 US-08-10-2 45 73.5 4.4 2 64 2 US-08-10-2 45 73.5 4.4 2 64 2 US-08-10-2 45 73.5 4.4 2 64 2 US-08-10-2 45 73.5 4.4 2 64 2 US-08-10-2 45 73.5 4.4 2 64 2 US-08-10-2 45 73.5 4.4 2 64 2 US-08-10-2 45 73.5 4.4 2 64 2 US-08-10-2 45 73.5 4.4 2 64 2 US-08-10-2 45 73.5 4.4 2 64 2 US-08-10-2 45 73.5 4.4 2 64 2 US-08-10-2 45 73.5 4.4 2 64 2 US-08-10-2 45 73.5 4.4 2 64 2 US-08-10-2 45 73.5 4.4 2 64 2 US-08-10-2 45 73.5 4.4 2 64 2 US-08-10-2 45 73.5 4.4 2 64 2 US-08-10-2 45 73.5 4.4 2 64 2 US-08-10-2 45 73.5 4.4 2 64 2 US-08-10-2 45 73.5 4.4 2 1 US-08-10-2 45 73.5 4.4 2 1 US-08-10-2 45 73.5 4.4 2 1 US-08-10-2 45 73.5 4.4 2 1 US-08-10-2 45 73.5 4.4 2 1 US-08-10-2 45 73.5 4.4 2 1 US-08-10-2 45 73.5 4.4 2 1 US-08-10-2 45 73.5 4.4 2 1 US-08-10-2 45 73.5 4.4 2 1 US-08-10-2 45 73.5 4.4 2 1 US-08-10-2 45 73.5 4.4 2 1 US-08-10-2 45 73.5 4.4 2 1 US-08-10-2 45 73.5 4.4 2 1 US-08-10-2 45 73.5 4.4 2 1 US-08-10-2 45 73.5 4.4 2 1 US-08-10-2 45 73.5 4.4 2 1 US-08-10-2 45 73.5 4.4 2 1 US-08-10-2 45 73.5 4.4 2 1 US-08-10-2 45 73.5 4 1 US-08-10-2 45 73.5 4 1 US-08-10-2 45 73.5 4 1 US-08-10-2 45 73.5 4 1 US-	Sequence 5, Appli Sequence 5, Appli Sequence 2, Appli Sequence 3, Appli Sequence 2, Appli Sequence 2, Appli Sequence 2, Appli Sequence 4, Appli Sequence 13, Appli Sequence 13, Appli Sequence 13, Appli Sequence 13, Appli Sequence 14, Appli Sequence 14, Appli Sequence 15, Appli Sequence 16, Appli Sequence 17, Appli Sequence 17, Appli Sequence 18, Appli Sequence 2, Appli Sequence 2, Appli Sequence 2, Appli Sequence 2, Appli Sequence 2, Appli Sequence 2, Appli	
75 4.5 704 1 75 4.5 704 1 75 4.5 704 1 74.5 4.5 369 3 74.5 4.5 369 3 74.5 4.5 369 3 74.5 4.5 381 1 74.5 4.5 4.5 381 1 74.5 4.5 4.2 3 74.5 4.5 4.2 3 74.5 4.5 4.2 3 74.5 4.5 4.2 3 74.5 4.5 242 3 74.5 4.5 263 4 74.4 5.5 263 4 74.4 6.5 263 6 700 TRY: Lawrencevile STRATION: WOMBER: US/O FILLING DATE: FILLIN	08 - 646 - 715 - 5 08 - 646 - 715 - 5 08 - 647 - 823 - 5 09 - 647 - 823 - 5 09 - 647 - 823 - 5 09 - 648 - 648 - 6 09 - 69 - 618 - 2 09 - 104 - 324 B - 6 09 - 104 - 324 B - 6 09 - 159 - 106 - 1 09 - 128 - 648 - 4	ALIGNMENTS 8 2el iles hoads lding 3, Suite 21 8,978 8,978
75 74.5 74.5 74.5 74.5 74.5 74.5 74.5 74	25.22 25.22 25.23 369 369 369 369 369 369 369 369 369 36	lication US/088 602 Brown, James R. Brown, James R. Brown, James R. Brown, James R. Dechert Price 97 Lenox Drive, renceville USA 3 DABLE FORM: E: Diskette B: Diskette Compatible SYSTEM: CSYSTEM: COMPATION: NUMBER: E: TION: A35 ATION DATA: N NUMBER: E: TION: A35 ATION DATA: N NUMBER: CSYSTEM: COM NUMBER: CSYSTEM: N NUMBER: CSYSTEM: COM NUMBER: CSYSTEM:
298 28 331 331 331 331 331 331 331 331 331 33	7	SGULT 1 108-898-978-2 SGQUENCE 2, APP PALENT NO. 6001 APPLICANT: APPLICANT: APPLICANT: APPLICANT: APPLICANT: APPLICANT: APPLICANT: APPLICANT: ADDRESSEE: OURRESPONDEN: ADDRESSEE: OURRESSEE: STRREET: ADDRESSEE: CONPUTER: CONPUTER: OPERATING SOGTWARE: OPERATING SOGTWARE: COMPUTER: OPERATING SOGTWARE: APPLICATION FILING DAT CLASSIFICAT ATTORNEY AREA APPLICATION FILING DAT ATTORNEY

Gaps

87;

42; Mismatches 104;

Ouery Match 5.6%; Score 93.5; DB 3; Best Local Similarity 19.7%; Pred. No. 0.39; Matches 57; Conservative 42; Mismatches 104;

Length 1041; Indels

```
qq
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     qq
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             Οy
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            ŏ
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    q
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             qq
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 QQ
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    δŏ
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  \overset{\circ}{\circ}
                                                    | : || || : : : | : : | : : | : : | : : | : : | : : | : : | : : | : : | : : | : : | : : | : : | : : | : : | : : | : : | : : | : : | : : | : : | : : | : : | : : | : : | : : | : : | : : | : : | : : | : : | : : | : : | : : | : : | : : | : : | : : | : : | : : | : : | : : | : : | : : | : : | : : | : : | : : | : : | : : | : : | : : | : : | : : | : : | : : | : : | : : | : : | : : | : : | : : | : : | : : | : : | : : | : : | : : | : : | : : | : : | : : | : : | : : | : : | : : | : : | : : | : : | : : | : : | : : | : : | : : | : : | : : | : : | : : | : : | : : | : : | : : | : : | : : | : : | : : | : : | : : | : : | : : | : : | : : | : : | : : | : : | : : | : : | : : | : : | : : | : : | : : | : : | : : | : : | : : | : : | : : | : : | : : | : : | : : | : : | : : | : : | : : | : : | : : | : : | : : | : : | : : | : : | : : | : : | : : | : : | : : | : : | : : | : : | : : | : : | : : | : : | : : | : : | : : | : : | : : | : : | : : | : : | : : | : : | : : | : : | : : | : : | : : | : : | : : | : : | : : | : : | : : | : : | : : | : : | : : | : : | : : | : : | : : | : : | : : | : : | : : | : : | : : | : : | : : | : : | : : | : : | : : | : : | : : | : : | : : | : : | : : | : : | : : | : : | : : | : : | : : | : : | : : | : : | : : | : : | : : | : : | : : | : : | : : | : : | : : | : : | : : | : : | : : | : : | : : | : : | : : | : : | : : | : : | : : | : : | : : | : : | : : | : : | : : | : : | : : | : : | : : | : : | : : | : : | : : | : : | : : | : : | : : | : : | : : | : : | : : | : : | : : | : : | : : | : : | : : | : : | : : | : : | : : | : : | : : | : : | : : | : : | : : | : : | : : | : : | : : | : : | : : | : : | : : | : : | : : | : : | : : | : : | : : | : : | : : | : : | : : | : : | : : | : : | : : | : : | : : | : : | : : | : : | : : | : : | : : | : : | : : | : : | : : | : : | : : | : : | : : | : : | : : | : : | : : | : : | : : | : : | : : | : : | : : | : : | : : | : : | : : | : : | : : | : : | : : | : : | : : | : : | : : | : : | : : | : : | : : : | : : | : : | : : : | : : | : : | : : | : : | : : | : : | : : | : : | : : | : : | : 
                                                                                                                                                                                                                          294 -LFPYFQDKKE-------LEAFRILPADFIEESEGTGIVHMAPAFGEADFF 336
                                                                                                                                                                                                                                                                                                                                   165 TAQFH-----TTEETL 198
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 199 QVCEQEDVPFVFDFHHFYANPDDHADLNVALPRMIKTWERIGLQPKVHLSSPKSEQAIRS 258
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                395 PFCWRTDSPLIYKAVNSW----FVAVEKVKSKMLKANESIHWTPE-HIKOGRFGKWLEG 448
112 NQFTLFTSPKESVTKNAVTDMAYHYRMLEAMGIADRSVIN-----IHIGGAYGNKDTA 164
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    194 YEKLSEKHSQNGFIVPPPPEKSLIGMTKVKVGKEDSSSAEFLEKRRAALERYLQRIV--- 250
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   77 THPDVMMDFVTPFQKEFREIGELVKTHQLRTSFHPNQFTLFTSPKESVTKNAV----TDM 132
                                                                                                                                                                                                                                                                                                                                                                                                              337 ACQEHNVPLVCPVDNQGCYTAEVKDFVGEYIKSADKGIARR--LKNENKLFYQGTVRHRY 394
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    ch 5.5%; Score 91.5; DB 1; Length 522; I Similarity 19.3%; Pred. No. 0.21; 64; Conservative 54; Mismatches 106; Indels 107; Gaps
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      28 YSKLSKTER-----KEALLTVTKANL-RNTMRTLHYIIGHGIPLYRFSSSIVPLA 76
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    APPLICANT: Gill, Gordon N.
APPLICANT: Kutten, Richard C.
APPLICANT: Cadena, Deborah L.
TITLE OF INVENTION: Sorting Nexins and Methods of Using Same NUMBER OF SEQUENCES: 25
CORRESPONDENCE ADDRESS:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               259 HADYVDANFLLERFRQWGTNIDFMIEAKQKDKALLRL--MDELSSIRGVK 306
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        --
--
--
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            COMPUTER REGABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM FC Compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentin Release #1.0, Version #1.25
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  ADDRESSEE: Campbell and Flores
STREET: 4370 La Jolla Village Drive, Suite 700
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        ATTORNEY/AGENT INFORMATION:
NAME: Campbell, Cathryn A.
REGISTRATION UNMARE: 31,815
REFERENCE/DOCKET NUMBER: P-UD 1955
TELECOMMUNICATION INFORMATION:
TELEPHONE: (619) 535-9001
TELEPHONE: (619) 535-9001
TELEPHONE: (519) 535-9001
SEQUENCE CHARACTERISTICS:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          APPLICATION NUMBER: US/08/625,322
FILING DATE: 01-APR-1996
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  Sequence 2, Application US/08625322
Patent No. 5804412
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   522 amino acids
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          CURRENT APPLICATION DATA:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         ; MOLECULE TYPE: protein US-08-625-322-2
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                San Diego
California
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   amino acid
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   linear
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 GENERAL INFORMATION:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      FILING DATE: 01 CLASSIFICATION:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    USA
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       92122
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          TOPOLOGY:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               US-08-625-322-2
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               COUNTRY:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                LENGTH:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                CITY:
STATE:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    Query Match
Best Local S
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       Matches
                                                                          qq
                                                                                                                                                             δŏ
                                                                                                                                                                                                                                                Op
                                                                                                                                                                                                                                                                                                                                   δ
                                                                                                                                                                                                                                                                                                                                                                                                         qq
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 δ
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                рp
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              δλ
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               Qy
```

```
186 ENDDKTYTTEETLQVCEQEDVPFVFDFHHFYANPD------DH----ADLNVALPRM 232
                                                                                                                                                                                                                       ::| | : | : | : | : | 356 SSEDNTALSRALSQLAEVEEK--:EQLHQEQANNDFFLLAELLSDYIRLLAIVRAAFDQR 413
                                                                                                                                                                                                                                                                                                                                                                -----WODAQATLQKKREAEARLLWANKPDKLQQAK 450
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   5.1%; Score 84.5; DB 4; Length 1073;
19.9%; Pred. No. 3.9;
tive 57; Mismatches 130; Indels 63; Gaps
251 NHPTMLQD---PDVREFLEKEELPRAVGTQTLSGAGLLKMFNKATDAVSKMTIKMNESDI 307
                                                         133 AYHYRMLEAMGIADR----SVINIHIGGAYGNKDTA--TAQFHQNIKQLPQEIKERMTL 185
                                                                                                                     308 WFEEKLQEVECEEORLRKLHAVVETLVN---HRKELALNTAQFAKSLAMLG------ 355
                                                                                                                                                                                                                                                                                                       233 IKTWERIGLQPKVHLSSPKSEQAIRSHADYVDANFLLERFRQ-----WGTNIDFMIEAK 286
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        688 -----SLQDISMSSQKLGNGISSELIELQKD-----MKESYRQLVQELRSLYNLQ----- 732
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  733 HTHEES-----QKEL------MYGVRND-----IDALVKTCTTSLNDADIILSDYI 772
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  571 YEKLDRIERNNKENENNFWNLKFNLLIMLRSFHGSFTDETNGYFTLLNDFNASMEELLNT 630
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 78 HPDVMWDFVTPPQKEFREIGELVKTHQLRTSFHPNQFTLFTSPKESVTKNAVTDMAYHYR 137
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    138 MLEAMGIADRSVINIHIGGAYGN-----KDTATAQFHQNIKQLPQEIKERMTLENDDKT 191
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         192 YTTEETLQVCEQEDVPFVFDFHHFYANPDDHADLNVALPRMIKTWERIGLQPKVHLSSPK 251
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            252 SEQAIRSHADYVD-----ANFLLERFRQWGTNIDFMIEAKQKDKALLRLMDELSSI 302
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               28 YSKLSKTER----KEALLTVTKANLRNTMRTLHYII-----GHGIPLYRFSSSIVPLA-T 77
                                                                                                                                                                                                                                                                                                                                                                                                                         287 QKDKALLRLMDELSSIRGVKRIGGGALQWKS 317
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  FastSEQ for Windows Version 4.0
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  451 -----LEWES 458
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            CURRENT APPLICATION NUMBER: US/09/541,782
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   Antifungal Assay
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    Sequence 6, Application US/09541782 Patent No. 6284480
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                ; Sequence 3, Application US/08947823 ; Patent No. 6114605
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         2000-04-03
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 APPLICANT: Nislow, Corey
APPLICANT: Sakowicz, Roman
APPLICANT: Baraud, Christophe
TITLE OF INVENTION: Antifungal
FILE REFERENCE: 1015
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      62; Conservative
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      CURRENT FILING DATE: 2000
NUMBER OF SEQ ID NOS: 10
                                                                                                                                                                                                                                                                                                                                                                414 MKTWQR-----
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 303 RGVKRIGGGALQ 314
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          833 RSEEFLRNAASO 844
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          Query Match
Best Local Similarity
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 GENERAL INFORMATION:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           ; TYPE: PRT
; ORGANISM: S.pombe
US-09-541-782-6
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              1073
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             US-08-947-823-3
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           US-09-541-782-6
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  SEQ ID NO 6
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  SOFTWARE:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              LENGIH:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   Matches
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   RESULT
```

```
US-08-886-886-11
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      qq
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                qq
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            Dp
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            QQ
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     οy
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  à
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            ŏ
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             16;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          75 LATHPDVMWDEVTPFQKEFREIGELVKTHQLRTSFHPN-----QFTLFTSPKESVTKNAV 129
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       259 LETSPDILREYLIPLOE-----HMVTVITPSTSGARNIHVMMEFLLL-----ILSDMP 306
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           130 TDMAYH---YRMLEAMGIADRSVINI-----HIGGAYGNKDT--ATAQFHQNIKQLPQE 178
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       | :| :|:|| :| :| :| 307 KDFIHHDKLFDLLDRVGVLTREVSTLVRDLEEEPRNKEGNNQTNCATLDLLENIELLKKD 366
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     179 IKERMTLENDDKTYTTEETLQVC-EQEDVPFVFDFHHFYANPDDHADLNVALPRMIKTWE 237
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      :| : |: | : | : | : | | 367 LK-HVYLKALDSS------QCCFPMSDGPLFMHLLHIHLN--DLLDSNAYSIALIK--E 414
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     238 RIGLQPKVHLSSPKSEQAIRSHADYVDANFLLERFRQWGTNIDFMIEAKQ-----KDK 290
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   59; Gaps
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           Length 1255;
GENERAL INFORMATION:
APPLICANT: Williamson, Valerie M.
APPLICANT: Kaloshian, Isgauhi
APPLICANT: Raghoobi, Jafar
APPLICANT: Paghoobi, Jafar
APPLICANT: Milligan, Stephen
TITLE OF INVENTION: Procedures and Materials for Conferring
TITLE OF INVENTION: Pest Resistance in Plants
NUMBER OF SEQUENCES: 5
CORRESPONDENCE ADDRESS:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           Query Match 5.1%; Score 84.5; DB 3; Length 1 Best Local Similarity 24.3%; Pred. No. 5.1; Matches 63; Conservative 37; Mismatches 100; Indels
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      SOFTWARE: PatentIn Release #1.0, Version #1.30
                                                                                                                                                                                                                                        E: Townsend and Townsend and Crew LLP
Two Embarcadero Center, Eighth Floor
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 NAME: Bastian, Kevin L. REGISTRATION NUMBER: 34,774
REFERENCE/DOCKET NUMBER: 023070-070210US
TELECOMMUNICATION INFORMATION:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  PRIOR APPLICATION DATA:
APPLICATION NUMBER: PCT/US97/18802
FILING DATE: 09-OCT-1997
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 60/028,191
FILING DATE: 10-OCT-1996
ATTORNEY/AGENT INFORMATION:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        CURRENT APPLICATION DATA:
APPLICATION NOMBER: US/08/947,823
FILING DATE: 09-0CT-1997
CLASSIFICATION: 800
                                                                                                                                                                                                                                                                                                                                                                                                              MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         TELEPAN: (415) 576-0200
TELEFAX: (415) 576-0300
INFORMATION FOR SEQ ID NO: 3: SEQUENCE CHARACTERISTICS: LENGTH: 1255 amino acids
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               | | |: | :|: | GLLHLIFSLPITIKKIKLI 483
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    291 ALLRLMDELS-SIRGVKRI 308
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                MOLECULE TYPE: protein
                                                                                                                                                                                                                                                                                                                                                                                       COMPUTER READABLE FORM: MEDIUM TYPE: Floppy
                                                                                                                                                                                                                                                                                        CITY: San Francisco
STATE: California
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   amino acid
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               Linear
                                                                                                                                                                                                                                                                                                                                                              94111-3834
                                                                                                                                                                                                                                                ADDRESSEE:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            TOPOLOGY:
                                                                                                                                                                                                                                                                                                                                          COUNTRY:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         US-08-947-823-3
                                                                                                                                                                                                                                                                    STREET:
CITY: SZ
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   465
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           òγ
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        ò
```

RESULT

```
12;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        83 WDFVTPFQKEFREIGEL-----VKTHQLRTSFHPNQFTLFTSPKESVTK-NAVTDMA 133
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     134 YHYRMLEAMGIADRSVIN-----IHIGGAYGNKDTATAQFHQNI-------KQL 175
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            128 FTVRNLEDTETIDTFITNTKAQRALVVGAGYISLEVLENLHHRGLDVTWIHRSTNINKLM 187
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           176 PQEIKERM--TLENDDKTYTTEETLQVCEQEDVPFVFDFHHFYANPDDHADLNVALPRMI 233
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          188 DQDMNQPIIDEIEKRNITYRFNEEISHVNGHEVTFT-----SGKVENFDLII----- 234
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              5.0%; Score 83.5; DB 3; Length 438;
19.6%; Pred. No. 1.2;
tive 46; Mismatches 115; Indels 69; Gaps
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       29 SKLSKTERKEALLTVTK-----ANLRNTMRTLHYIIGHGIPLYRFSSSIVPLATHPDVM 82
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       18 SQIRRLDKESEIIVFEKDRDMSFANC----ALPYYIGNVIEDRR-----KVLAYTPNOF 67
                                                                                                                                                                             AND INHIBITORS THEREOF USEFUL AS ANTIMICROBIAL
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            234 KTWERIGLQPKVHLSSPKSEQAIRSHADYVDANFLLERFRQWGTNI 279
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           235 ---EGVGTHPNSQFI--KSSNVILNDKGYIPVN-----HNFQTNI 269
                                                                    APPLICANT: Ratz, Leonard
APPLICANT: Bales and Stephen B.
APPLICANT: Delcardayre, Stephen B.
APPLICANT: Delcardayre, Stephen B.
TITLE OF INVENTION: COENZYME A DISULFIDE REDUCTASE,
TITLE OF INVENTION: AND INHIBITORS THEREOF USEFUL AS
TITLE OF SEQUENCES: 40
CORRESPONDENCE ADDRESS:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        OPERATING SYSTEM: DOS
SOFTWARE: FastSEG for Windows Version 2.0
CURRENT APPLICATION DATA:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      6016.US.P1
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          UMBER: US/08/886,886
02-JUL-1997
; Sequence 11, Application US/08886886
; Patent No. 6107068
                                                                                                                                                                                                                                                                                 ADDRESSEE: Abbott Laboratories
STREET: 100 Abbott Park Road
CITY: Abbott Park
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    NAME: Casuto, Dianne
REGISTRATION NUMBER: 40,943
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           TELECOMMUNICATION INFORMATION:
TELEPHONE: 847-938-3137
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      COMPUTER: IBM Compatible OPERATING SYSTEM: DOS
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             ATTORNEY/AGENT INFORMATION:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      REFERENCE/DOCKET NUMBER:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    438 amino acids
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              Best Local Similarity 19.6
Matches 56; Conservative
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               INFORMATION FOR SEQ ID NO: SEQUENCE CHARACTERISTICS: LENGTH: 438 amino acid
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             TELEFAX: 847-938-2623
                                                                                                                                                                                                                                                                                                                                                                                                                                                                 Diskette
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         CLASSIFICATION: 435
PRIOR APPLICATION DATA:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   single
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                MOLECULE TYPE: protein
                                                                                                                                                                                                                                                                                                                                                                                                                                 COMPUTER READABLE FORM:
MEDIUM TYPE: Diskett
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       APPLICATION NUMBER: FILING DATE: 02-JUI
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             APPLICATION NUMBER:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           unknown
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              amino acid
                                                                                                                                                                                                                                                                                                                                                                                                          60064-3500
                                                       GENERAL INFORMATION:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   STRANDEDNESS:
                                                                                                                                                                                                                                                                                                                                                                                     USA
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     FILING DATE:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      TELEPHONE:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                TOPOLOGY:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             US-08-886-886-11
                                                                                                                                                                                                                                                                                                                                                                                  COUNTRY:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           Query Match
```

```
120 PKESVTKNAVTDMAYHYRMLEAMGIADRSVINIHIGGA---YGN------- 160
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              161 -----KDTATAQFHQ----NIKQLPQEIKERMTLENDDKTYTTEETLQVCEQED 205
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                79 GWKKKKEQMERDGCSEQESQPCAFIGIGNSDQEMQQ---LNLEGKNYCTAKTLYISDSDK 135
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      206 -VPFVFDFHHFYANPDDHADLNVALPRMIKTWERIGLQPKVHLSSP-KSEQAIRSHADYV 263
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   264 DANFLLERFRQWGTNIDFMIEAKQKDKALLRLMDELSSIRGVKRIGG----GALQW 315
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    176 ------NADLCIASGTKVALFNRLRSQTVSTRYLHVEGGNFHASSQQW 217
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 85;
                                                                                  APPLICANT: Peterson, Michael G
APPLICANT: Henkel, Thomas
TITLE OF INVENTION: EPSTEIN-BARR VIRUS TRANSCRIPTION FACTOR
TITLE OF INVENTION: BINDING ASSAY
NUMBER OF SEQUENCES: 11
CORRESPONDENCE ADDRESS:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             Ouery Match 5.0%; Score 82.5; DB 1; Length 500; Best Local Similarity 21.6%; Pred. No. 1.9; Matches 51; Conservative 32; Mismatches 68; Indels 8:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          EPSTEIN-BARR VIRUS TRANSCRIPTION FACTOR
                                                                                                                                                                                                                            ADDRESSEE: FLEHR, HOHBACH, TEST, ALBRITTON & HERBERT STREET: 4 Embarcadero Center, Suite 3400
                                                                                                                                                                                                                                                                                                                                                                                                                                                   SOFTWARE: Patentin Release #1.0, Version #1.25 CURRENT APPLICATION DATA:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           REFERENCE TOCKET NUMBER: A 59233/RAO
TELECOMMUNICATION INFORMATION:
TELEPHONE: (415) 781-1989
TELERAX: (415) 398-3249
TELEX: 910 277299
INFORMATION FOR SEQ ID NO: 2:
SEQUENCE CHARACTERISTICS:
LENGTH: 500 amino acids
TYPE: amino acid
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              PRIOR APPLICATION DATA:
APPLICATION NUMBER: US/08/246,977
FILING DATE: 20-MAY-1994
ATTORNEY/AGENT INFORMATION:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              APPLICATION NUMBER: US/08/704,398
                                                                                                                                                                                                                                                                                                                                                                                                                                PC-DOS/MS-DOS
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        Sequence 2, Application PC/TUS9505966 GENERAL INFORMATION:
                       Sequence 2, Application US/08704398
Patent No. 5679525
                                                                                                                                                                                                                                                                                                                                                                                   E: Floppy disk
IBM PC compatible
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   NAME: Osman, Richard A
REGISTRATION NUMBER: 36,627
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  APPLICANT: TULARIK, INC. TITLE OF INVENTION: EPST
                                                                                                                                                                                                                                                                                                                                                         COMPUTER READABLE FORM: MEDIUM TYPE: Floppy
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        MOLECULE TYPE: protein
                                                                                                                                                                                                                                        STREET: 4 Embarracisco
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         CLASSIFICATION: 435
                                                                                                                                                                                                                                                                                                                                                                                                      COMPUTER: IBM PC OPERATING SYSTEM:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       linear
                                                                    GENERAL INFORMATION:
                                                                                                                                                                                                                                                                                                                                      94111-4187
                                                                                                                                                                                                                                                                                                                      USA
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   FILING DATE
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     TOPOLOGY:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  PCT-US95-05966-2
US-08-704-398-2
                                                                                                                                                                                                                                                                                                                 COUNTRY:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             US-08-704-398-2
                                               Patent No.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             RESULT
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 q
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 qq
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    δλ
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              ò
```

```
68; Indels 85; Gaps
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     206 -VPFVFDFHHFYANPDDHADLNVALPRMIKTWERIGLQPKVHLSSP-KSEQAIRSHADYV 263
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        161 ------KDTATAQFHQ-----NIKQLPQEIKERMTLENDDKTYTTEETLQVCEQED 205
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          27 PPKRLTREAMRN-----YLKERG--DQTVLILHAKVAQKSYGNEKRFFCPPPCVYLMGS 78
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              264 DANFLLERFRQWGTNIDFMIEAKQKDKALLRLMDELSSIRGVKRIGG----GALQW 315
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         176 ------NADLCIASGTKVALFNRLRSQTVSTRYLHVEGGNFHASSQQW 217
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          5.0%; Score 82.5; DB 5; Length 500; 21.6%; Pred. No. 1.9; Live 32; Mismatches 68; Indels 86
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    APPLICANT: STAFFÖRD, DOUGLAS C.
TITLE OF INVENTION: VACCINE AND ANTITOXIN FOR TREATMENT AND
TITLE OF INVENTION: PREVENTION OF C. DIFFICILE DISEASE
NUMBER OF SEQUENCES: 32
                AUVESSEE: FLEHR, HOUBACH, TEST, ALBRITTON & HERBERT STREET: 4 Embarcadero Center, Suite 3400 CITY: San Francisco STATE: Californi
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               120 PKESVTKNAVTDMAYHYRMLEAMGIADRSVINIHIGGA---YGN-----
                                                                                                                                                                                                                                                                                                SOFTWARE: Patentin Release #1.0, Version #1.25 CURRENT APPLICATION DATA:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    REFERENCE/DOCKET NUMBER: FP-59233-PC/RAO TELECOMMUNICATION INFORMATION:
                                                                                                                                                                                                                                                                                                                                              APPLICATION NUMBER: PCT/US95/05966
                                                                                                                                                                                                                                                                                                                                                                                                          PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/246,977
FILING DATE: 20-MAX-1994
ATTORNEY/AGENT INFORMATION:
                                                                                                                                                                                                                                                  COMPUTER: IBM PC compatible OPERATING SYSTEM: PC-DOS/MS-DOS
TITLE OF INVENTION: BINDING ASSAY
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         Sequence 6, Application US/08480604A Patent No. 5736139
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              36,627
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            APPLICANT: KINK, JOHN A.
APPLICANT: THALLEY, BROCE S.
APPLICANP: PADHYE, NISHA V.
APPLICANT: FIRCA, JOSEPH R.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  (415) 781-1989
(415) 398-3249
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 INFORMATION FOR SEQ ID NO: 2: SEQUENCE CHARACTERISTICS:
                                                                                                                                                                                                                                 Floppy disk
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 500 amino acids
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                Query Match
Best Local Similarity 21.6%
Matches 51; Conservative
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        NAME: Osman, Richard A REGISTRATION NUMBER: 36
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                MOLECULE TYPE: protein
                                                                                                                                                                                                        COMPUTER READABLE FORM:
                     NUMBER OF SEQUENCES: 1
CORRESPONDENCE ADDRESS:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       NUMBER OF SEQUENCES: 32
CORRESPONDENCE ADDRESS:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              TELEX: 910 277299
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       amino acid
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               linear
                                                                                                                                                                                94111-4187
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       GENERAL INFORMATION:
                                                                                                                                                                                                                                                                                                                                                                                            CLASSIFICATION:
                                                                                                                                                               USA
                                                                                                                                                                                                                            MEDIUM TYPE:
                                                                                                                                                                                                                                                                                                                                                                      FILING DATE:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  TELEPHONE:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           TOPOLOGY:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      PCT-US95-05966-2
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   US-08-480-604A-6
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          TELEFAX:
                                                                                                                                                             COUNTRY:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 LENGTH:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             RESULT
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    ŏλ
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              δ
```

ADDRESSEE: MEDLEN & CARROLL, LLP

QQ á

```
88 PFQKEFREIGELVKTHQLRTSFHPNQFTLFTSPK---ESVTKNAV------TDM 132
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        169 KFYKKRMEF -- IYDRQKRFINYYKSQINKPTVPTIDDIIKSHLVSEYNRDETVLESYRTN 226
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     ----- DVPFVFDFHHF---Y 216
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        227 SL----RKINSNHGIDIRANSLFTEQELLNIYSQELLNRGNLAAASDIVRLLALKNFGGVY 283
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    AN----PDDHADLNVALPRM----IKTWERIGLQPKVHLSSPKSEQAIRSHADYVDANFL 268
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         Gaps
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       133 AYHYRMLEAMGIADR--------SVINIHIGGAYGNKDTATAQFHQN 171
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            29 SKLSKTERKEALLTVTKANLRNTMRTLHYI-IGHGIPLYRFSSSIVPLATHPDVMWDFVT 87
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         Indels 105;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         5.0%; Score 82.5; DB 1; Length 2710; 18.2%; Pred. No. 29; tive 54; Mismatches 116; Indels 105;
                                                                                                                                                                                                PatentIn Release #1.0, Version #1.30
STREET: 220 MONTGOMERY STREET, SUITE 2200 CITY: SAN FRANCISCO STATE: CALIFORNIA
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   172 IKQLPQEIKERMTLE-NDDKTYTTEETLQVCEQE-
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         NAME: INGOLIA, DIANE E.
REGISTATION NUMBER: 40,027
REFERENCE/DOCKET NUMBER: 0PHD-01763
TELECOMMUNICATION INFORMATION:
TELEPRONE: (415) 705-8410
TELERAX: (415) 397-8338
INFORMATION FOR SEQ ID NO: 6:
                                                                                                                                                                                                                                            APPLICATION NUMBER: US/08/480,604A FILING DATE: 07-JUN-1995
                                                                                                                                                                                                                                                                                                                                                                                     APPLICATION NUMBER: US 08/405,496 FILING DATE: 16-MAR-1995 PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/329,154 FILING DATE: 25-OCT-1994 PRIOR APPLICATION DATA:
                                                                                                                                                                                                                                                                                                                        UNITED STATES OF AMERICA
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      US 08/161,907
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               APPLICATION NUMBER: US 07/985,321
FILING DATE: 04-DEC-1992
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 07/429,791
FILING DATE: 31-OCT-1989
ATTORNEY/AGENT INFORMATION:
                                                                                                             COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                APPLICATION NUMBER: US 08 FILING DATE: 02-DEC-1993 PRIOR APPLICATION DATA:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        SEQUENCE CHARACTERISTICS:
LENGTH: 2710 amino acids
                                                                                                                                                                                                                    CURRENT APPLICATION DATA:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   61; Conservative
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  ; MOLECULE TYPE: protein US-08-480-604A-6
                                                                                                                                                                                                                                                                                                       PRIOR APPLICATION DATA:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     amino acid
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             linear
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               Best Local Similarity
Matches 61; Conserv
                                                                                                                                                                                                                                                                                 CLASSIFICATION:
                                                                                           94104
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       TOPOLOGY:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              Query Match
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       a
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 ŏ
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        a
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                ò
```

```
17;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     72 SNLKKDILKEVIL-IKNSNTSPVEKNLHFVWIGGEVS------DIAL 111
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          112 EYIKQWADIN---AEYNIKLWYDSEAFLVNTLKKAIVESSTTEALQLLEEEIQNPQFDNM 168
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   133 AYHYRMLEAMGIADR------SVINIHIGGAYGNKDTATAQFHQN 171
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        169 KFYKKRMEF--IYDRQKRFINYYKSQINKPTVPTIDDIIKSHLVSEYNRDETVLESYRTN 226
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            29 SKLSKTERKEALLTVTKANLRNTMRTLHYI-IGHGIPLYRFSSSIVPLATHPDVMWDFVT 87
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                5.0%; Score 82.5; DB 2; Length 2710;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       Indels
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        88 PFQKEFREIGELVKTHQLRTSFHPNQFTLFTSPK---ESVTKNAV----
                                                                                             TITLE OF INVENTION: VACCINE FOR CLOSTRIDIUM BOTULINUM IITLE OF INVENTION: NEUROTOXIN
                                                                                                                                                                                                                                                                                                                                                     54; Mismatches 116;
                                                                                                                                                                           E: MEDLEN & CARROLL, LLP
220 MONTGOMERY STREET, SUITE 2200
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     Pred. No. 29;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       REFERENCE/DOCKET NUMBER: OPHD-01308 TELECOMMUNICATION INFORMATION:
                                                                                                                                                                                                                                                                                                                                                                                                          J6-MAR-1995
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    FILING DATE: 04-DEC-1992
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 07/400 TO FILING DATE:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     APPLICATION NUMBER: US 07/429,791
FILING DATE: 31-OCT-1989
ATTORNEY/AGENT INFORMATION:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         US 08/161,907
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                APPLICATION NUMBER: US 08/329,154
                                                                                                                                                                                                                                                                                                                MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
                      Sequence 6, Application US/08405496A
Patent No. 5919665
                                                                             WILLIAMS, JAMES A.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       40,027
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              TELEPHONE: (415) 705-8410
TELEFAX: (415) 397-8338
INFORMATION FOR SEQ ID NO: 6:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              FILING DATE: 25-OCT-1994
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           02-DEC-1993
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        2710 amino acids
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 18.2%;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  NAME: INGOLIA, DIANE E. REGISTRATION NUMBER: 4(
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     Conservative
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         SEQUENCE CHARACTERISTICS:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              PRIOR APPLICATION DATA: APPLICATION NUMBER: 1
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 ; MOLECULE TYPE: protein US-08-405-496A-6
                                                                                                                                                                                                                                                                                                                                                                                                                                                                           PRIOR APPLICATION DATA:
                                                                                                                                                           CORRESPONDENCE ADDRESS:
                                                                                                                                                                                                                                                                                              COMPUTER READABLE FORM:
                                                                                                                                                                                                                CITY: SAN FRANCISCO
STATE: CALIFORNIA
COUNTRY: USA
                                                                                                                                      NUMBER OF SEQUENCES:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                amino acid
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  linear
                                                         GENERAL INFORMATION:
APPLICANT: WILLIAN
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             Best Local Similarity
Matches 61; Conserv
                                                                                                                                                                                                                                                                                                                                                                                                                                                         CLASSIFICATION:
                                                                                                                                                                                                                                                                                                                                                                                                                                      FILING DATE:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             FILING DATE:
                                                                                                                                                                                                                                                                           94.104
                                                                                                                                                                           ADDRESSEE:
US-08-405-496A-6
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            TOPOLOGY:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            LENGIH:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                Query Match
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              Οý
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               Dp
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            Op
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   οy
```

⋩	172 IKQLPQEIKERMTLE-NDDKTYTTEETLQVCEQEDVPFVFDFHHF'Y 216 :: ::	
ą	227 SLRKINSNHGIDIRANSLFTEQELLNIYSQELLNRGNLAAASDIVRLLALKNFGGVY 283	
λ.	217 ANPDDHADLNVALPRMIKTWERIGLOPKVHLSSPKSEQAIRSHADYVDANFL 268	
ā	LDVDMLKGIRSDLFNIISKKSSIGLDRWEMINL************************************	
± 9	269 LERFROWGTNIDFMTEAKOKDKALLKIMDELS 300 	
RESI JS - St	RESULT 10 15-08-915-136-6 : Sequence 6, Application US/08915136	
Δ,	D. 6290960 INFORMATION:	
	ANT:	
	PADHYE, NISHA V FIRCA, JOSEPH R	
	D) DOUG	
	CENTRALON: PREVENTION OF C. DIFFICILE DISEASE CENTRALON: 23	
	E ADDRESS:	
	n :-	
	CITY: SAN FRANCISCO STATE: CALIFORNIA	
	Y 0	
	ER READABLE F	
	MEDIUM TYPE: Floppy disk COMPUTER: IBM PC compatible	
	OPERATING SYSTEM: PC-DOS/MS-DOS	
	** ", verston **	
	APPLICATION NUMBER: US/08/915,136	
	CLASSIFICATION:	
	PRIOR APPLICATION DATA:	
	NOMBER:	
	TION DATA:	
	APPLICATION NUMBER: US 08/405,496 FILING DATE: 16-MAR-1995	
	A:	
	APPLICATION NUMBER: US 08/329,154	
	A:	
	APPLICATION NUMBER: US 08/161,907	
	FILING DAIE: 02-DEC 1993 PRIOR APPLICATION DATA:	
	APPLICATION NUMBER: US 07/985,321	
	FILING DATE: 04-DEC-1992 PRIOR APPLICATION DATA:	
	APPLICATION NUMBER: US 07/429,791	
., .	FILING DATE: 31-OCT-1989	
	- ⊡	
	027	
	REFERENCE/DOCKET NUMBER: OPHD-U1/63 TELECOMMUNICATION INFORMATION:	
	E: (415	
	TELEFAX: (415) 39/-8338 INFORMATION FOR SEQ ID NO: 6:	
	HARACTERISTI	
	LENGTH: 2/10 dmino acid	
	TOPOLOGY: linear	

```
9
                                                                                                                                        17;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            172 IKQLPQEIKERMTLE-NDDKTYTTEETLQVCEQE------DVPFVFDFHHF---Y 216
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 227 SL---RKINSNHGIDIRANSEFTEQELENIYSQEILINRGNEAAASDIVRLEAEKNFGGVY 283
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    217 AN----PDDHADLNVALPRM----IKTWERIGLQPKVHLSSPKSEQAIRSHADYVDANFL 268
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                284 LDVDMLPGIHSDLFKTISRPSSIGLDRWEMIKL------BAIMKYKKXIN-NYT 330
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                213 HHFYANPDDHADLNVALPRMIKTWERIGLQPK----VHLSSP-----KSEQAIRSHA 260
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             261 DYVDANFLLER-----FRQWGTNIDFMI--EAKQKDKALLRLMDELSSIRGVKRIGGG 311
                                                                                                                                                                                                                   72 SNLKKDILKEVIL-1KNSNTSPVEKNLHFVWIGGEVS-------DIAL 111
                                                                                                                                                                                                                                                                                        88 PEQKEFREIGELVKTHQLRTSFHPNQFTLFTSPK---ESVTKNAV-----TDM 132
                                                                                                                                                                                                                                                                                                                                  112 EYIKQWADIN---AEYNIKLWYDSEAFLVNTLKKAIVESSTTEALQLLEEEIQNPOFDNM 168
                                                                                                                                                                                                                                                                                                                                                                                     ---SVINIHIGGAYGNKDTATAQFHQN 171
                                                                                                                                                                                                                                                                                                                                                                                                                                   169 KFYKKRMEF -- IYDRQKRFINYYKSQINKPTVPTIDDIIKSHLVSEYNRDETVLESYRTN 226
                                                                                                                                           Gaps
                                                                                                                                                                                      29 SKLSKTERKEALLTVTKANLRNTMRTLHYI-IGHGIPLYRFSSSIVPLATHPDVMWDFVT 87
                                                                                     Query Match 5.0%; Score 82.5; DB 4; Length 2710; Best Local Similarity 18.2%; Pred. No. 29; Matches 61; Conservative 54; Mismatches 116; Indels 105;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              Ouery Match 4.9%; Score 82; DB 6; Length 323; Best Local Similarity 26.2%; Pred. No. 1.1; Matches 32; Conservative 18; Mismatches 48; Indels 24;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  Sequence 12, Application US/07977434
Patent No. 5466591
GENERAL INFORMATION:
APPLICANT: Gelfand, David H.
APPLICANT: Abramson, Richard D.
TITLE OF INVENTION: 5, TO 3' EXONUCLEASE MUTATIONS OF TITLE OF INVENTION: THERMOSTABLE DNA POLYMERASES
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   269 LERF----RQWGTNIDFMIEAKQKDKALLRLMDELS 300
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            331 SENFDKLDQQLKDNFKLIIESKSEKSEIFSKLENLN 366
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               APPLICANT: LINNENBACH, ALBAN
                                                                                                                                                                                                                                                                                                                                                                                     133 AYHYRMLEAMGIADR-----
MOLECULE TYPE: protein
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        ; Patent No. 5185254
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    LENGTH: 323
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     RESULT 12
US-07-977-434-12
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 312 AL 313
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             277 LI 278
                       US-08-915-136-6
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 5185254-4
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         5185254-4
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        RESULT
                                                                                                                                                                                                                                                                                                                                                                                                                                      Op
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          ŏλ
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   qq
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      Qy
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   δy
                                                                                                                                                                                                                                                                                        Q \underline{y}
                                                                                                                                                                                                                                                                                                                                                                                     \Omega
```

```
534 EKVFEIAGETENLNSSTQVA--YILFEKLNIAP------YKKTATGKFSTNAE 578
                                                                                               414 LNPDEKRFNLEELSLKYLGYKMISFDELVNENVPLFGNDFSYVPLERAVEYSCEDADVTY 473
                                                                                                                                                                                                                                                                                                  117 ---FTSPKESVTKNAVTDMAYHYRMLEAMGIADRSVINTHIGGAYGNKDTATAQFHQNIK 173
                                                                                                                                                                                                                                                                                                                                                                                                                                                                         | : || : || : || 579 VLEELSKEHEIAKLLLE-----YRKYQKLKSTYIDSIPLSINRKTNRVHTTFHQTGTSTG 633
                                                                                                                                                                                                                          474 RIFRKLGRKIYENEMEKLFYEIEMPLIDVLSEMELNGVYFDEEYLKELSKKYQEKMDGIK 533
                                                                                                                                                                                                                                                                                                                                                                                                                                174 QLPQEIKE----FHH----- 214
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  --DDHADLNVALPRMIKTW-----ERIGLQPKVHLSSPKSEQAI 256
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            634 RLSSSNPNLQNLPTRSEEGKEIRKAVRPQRQDWWILGADYSQIELRVLAHVS--KDENLL 691
                                  31 LSKTERKEALLTVTKANLRNTMRTLHYIIGHGIPLYRFSSSIVPLATHPDVMWDFVTPFQ 90
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     GENERAL INFORMATION:
GENERAL INFORMATION:
GENERAL APPLICANT: Gelfand, David H.
APPLICANT: Abramson, Richard D.
TITLE OF INVENTION: 5' TO 3' EXONUCLEASE MUTATIONS OF
TITLE OF INVENTION: THERMOSTABLE DNA POLYMERASES
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             257 RSHADYVDAN------FLLERFRQWGTNIDFMI 283
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   JMBER: US/08/458,819
02-JUN-1995
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          Hoffmann-La Roche Inc.
340 Kingsland Street
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               UMBER: US 590,490
28-SEP-1990
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            UMBER: US 590,466
28-SEP-1990
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            APPLICATION NUMBER: US 590,213 FILING DATE: 28-SEP-1990
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              JMBER: US 063,509
17-JUN-1987
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        Sequence 12, Application US/08458819
Patent No. 5795762
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            07/977,434
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               JMBER: US 523,394
15-MAY-1990
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               APPLICATION NUMBER: US 143,441
                                                                                                                                                                91 KEFREIGELVKTHQLRTSFHPNQFTL --
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       OPERATING SYSTEM: 7
SOFTWARE: WordPerfect 2.1
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08,
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          12-JAN-1988
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       FILING DATE: 23-FEB-1993
APPLICATION NUMBER: US 5
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   38
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         PRIOR APPLICATION DATA: APPLICATION NUMBER: (
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           FILING DATE: 28-SEP-
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             FILING DATE: 12-JAN-:
PRIOR APPLICATION DATA:
APPLICATION NUMBER: U
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          PRIOR APPLICATION DATA:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             PRIOR APPLICATION DATA:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         PRIOR APPLICATION DATA:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            NUMBER OF SEQUENCES: 3
CORRESPONDENCE ADDRESS:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           Macintosh
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            APPLICATION NUMBER:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               APPLICATION NUMBER:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                : New Jersey
07110-1199
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           CLASSIFICATION:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              Nutley
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           FILING DATE:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               FILING DATE:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              FILING DATE:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               215 --FYANP-
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             ADDRESSEE:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              RESULT 13
US-08-458-819-12
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               COMPUTER:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   STREET:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             STATE:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                QQ
                                                                                               ^{\mathrm{q}}
                                                                                                                                                                                                                                                                                                                                                               q
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              q
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               Qy
                                                                                                                                                                   ά
                                                                                                                                                                                                                                 qq
                                                                                                                                                                                                                                                                                                  δ
                                                                                                                                                                                                                                                                                                                                                                                                                                    ŏ
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       REFERENCE/DOCKET NUMBER: Case No. 5466591 8753
TELECOMMUNICATION INFORMATION:
TELEPHONE: (510) 814-2972
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                FILING DATE: 21-DEC-1990
PRIOR APPLICATION DATA:
APPLICATION AND APPLICATION DATA:
                                                                                                                                                                                                                                                                                                                                                     SOFTWARE: WordPerfect 2.1
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/07/977,434
                                                             3: Hoffmann-La Roche Inc.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    APPLICATION NUMBER: US 590,466 FILING DATE: 28-SEP-1990 APPLICATION DATA: APPLICATION WUMBER: US 590,213 FILING DATE: 28-SEP-1990
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   APPLICATION NUMBER: US 143,441 FILING DATE: 12-JAN-1988 APPLICATION DATA: APPLICATION NUMBER: US 063,509 FILING DATE: 17-JUN-1987
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      APPLICATION NUMBER: US 899,241
FILING DATE: 22-AUG-1986
PRIOR APPLICATION DATA:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 590,490
FILING DATE: 28-5EP-1990
PRIOR APPLICATION DATA:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       APPLICATION NUMBER: US 523,394
FILING DATE: 15-MAY-1990
PRIOR APPLICATION DATA:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         APPLICATION NUMBER: US 746,121 FILING DATE: 15-AUG-1991
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     APPLICATION NUMBER: US 585,471 FILING DATE: 20-SEP-1990 PRIOR APPLICATION DATA:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 609,157
FILING DATE: 02-NOV-1990
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       US 557,517
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        US 455,611
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     REGISTRATION NUMBER: 31,822
                                                                                                                                                                                                                          COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: Macintosh
OPERATING SYSTEM: 7
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        APPLICATION NUMBER: US 45
FILING DATE: 22-DEC-1989
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      FILING DATE: 24-JUL-1990 ATTORNEY/AGENT INFORMATION:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        (510) 814-2972
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        : 892 amino acids
amino acid
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         SEQUENCE CHARACTERISTICS:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       INFORMATION FOR SEQ ID NO:
NUMBER OF SEQUENCES: 38
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         FILING DATE: 02-NOV-1
PRIOR APPLICATION DATA:
APPLICATION NUMBER: U
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    ; MOLECULE TYPE: protein US-07-977-434-12
                                  CORRESPONDENCE ADDRESS:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        PRIOR APPLICATION DATA:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       PRIOR APPLICATION DATA:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      PRIOR APPLICATION DATA:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       CLASSIFICATION: 435
                                                                                                                                                             STATE: New Jersey ZIP: 07110-1199
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             Luann Cserr
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            linear
                                                                       STREET: STREET STREET STREET STREET STREET STREET STREET STREET STREET STREET STREET STREET STREET STREET STREET STREET STREET STREET STREET STREET STREET STREET STREET STREET STREET STREET STREET STREET STREET STREET STREET STREET STREET STREET STREET STREET STREET STREET STREET STREET STREET STREET STREET STREET STREET STREET STREET STREET STREET STREET STREET STREET STREET STREET STREET STREET STREET STREET STREET STREET STREET STREET STREET STREET STREET STREET STREET STREET STREET STREET STREET STREET STREET STREET STREET STREET STREET STREET STREET STREET STREET STREET STREET STREET STREET STREET STREET STREET STREET STREET STREET STREET STREET STREET STREET STREET STREET STREET STREET STREET STREET STREET STREET STREET STREET STREET STREET STREET STREET STREET STREET STREET STREET STREET STREET STREET STREET STREET STREET STREET STREET STREET STREET STREET STREET STREET STREET STREET STREET STREET STREET STREET STREET STREET STREET STREET STREET STREET STREET STREET STREET STREET STREET STREET STREET STREET STREET STREET STREET STREET STREET STREET STREET STREET STREET STREET STREET STREET STREET STREET STREET STREET STREET STREET STREET STREET STREET STREET STREET STREET STREET STREET STREET STREET STREET STREET STREET STREET STREET STREET STREET STREET STREET STREET STREET STREET STREET STREET STREET STREET STREET STREET STREET STREET STREET STREET STREET STREET STREET STREET STREET STREET STREET STREET STREET STREET STREET STREET STREET STREET STREET STREET STREET STREET STREET STREET STREET STREET STREET STREET STREET STREET STREET STREET STREET STREET STREET STREET STREET STREET STREET STREET STREET STREET STREET STREET STREET STREET STREET STREET STREET STREET STREET STREET STREET STREET STREET STREET STREET STREET STREET STREET STREET STREET STREET STREET STREET STREET STREET STREET STREET STREET STREET STREET STREET STREET STREET STREET STREET STREET STREET STREET STREET STREET STREET STREET STREET STREET STREET STREET STREET STREET STREET STREET STREET STREET STREET STREET STREET STREET STREET S
                                                                                                                                                                                                                                                                                                                                                                                                                                                               FILING DATE:
                                                                ADDRESSEE:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         TOPOLOGY:
```

US 899,241

PRIOR APPLICATION DATA:
APPLICATION NUMBER:

11;

Query Match

4.9%; Score 81.5; DB 1; Length 892;
Best Local Similarity 17.1%; Pred. No. 6.2;
Matches 58; Conservative 57; Mismatches 116; Indels 109; Gaps

FILING DATE:

```
31 LSKTERKEALLTVTKANLRNTMRTLHY11GHG1PLYRFSSSIVPLATHPDVMWDFVTPFQ 90
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          TOPOLOGY:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                LENGTH:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            Query Match
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         Matches
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    534 EKVFEIAGETFNLNSSTQVA--YILFEKLNIAP-----YKKTATGKFSTNAE 578
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  474 RIFRKLGRKIYENEMEKLFYEIEMPLIDVLSEMELNGVYFDEEYLKELSKKYQEKMDGIK 533
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      117 ---FTSPKESVTKNAVTDMAYHYRMLEAMGIADRSVINIHIGGAYGNKDTATAQFHQNIK 173
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  174 QLPQEIKE----FMTLENDDKTYTTEETLQVCEQEDVPFVFD------FHH----- 214
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             215 --FYANP-----DDHADLNVALPRMIKTW------ERIGLQPKVHLSSPKSEQAI 256
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         414 LNPDEKRFNLEELSLKYLGYKMISFDELVNENVPLFGNDFSYVPLERAVEYSCEDADVTY 473
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             31 LSKTERKEALLIVIKANLRNIMRTLHYIIGHGIPLYRFSSSIVPLATHPDVMWDFVIPFQ 90
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     57; Mismatches 116; Indels 109;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   Query Match 4.9%; Score 81.5; DB 1; Length 892; Best Local Similarity 17.1%; Pred. No. 6.2;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            APPLICANT: Gelfand, David H.
APPLICANT: Gelfand, David H.
APPLICANT: Abramson, Richard D.
TITLE OF INVENTION: 5' TO 3' EXONUCLEASE MUTATIONS OF
TITLE OF INVENTION: THERMOSTABLE DNA POLYMERASES
CORRESPONDENCE ADDRESS:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      REFERENCE/DOCKET NUMBER: Case No. 5795762 8753 TELECOMMUNICATION INFORMATION:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         257 RSHADYVDAN------FLLERFRQWGTNIDFMI 283
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     692 KAFKEDLDIHTITAAKIFGVSEMFVSEQMRRVGKMVNFAI 731
                                                                                                               APPLICATION NUMBER: WO PCT/US90/07641 FILLING DATE: 21-DEC-1990
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          91 KEFREIGELVKTHQLRTSFHPNQFTL-----
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  Sequence 12, Application PC/TUS9107035 GENERAL INFORMATION:
                                                                                                                                                                                         APPLICATION
FILING DATE: 20-SEF-12.2
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 455,611
                                                                                                                    FILING DATE: 02-NOV-1990
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 0S 557,517
FILING DATE: 24-JUL-1990
ATTORNEY/AGENT INFORMATION:
                                          APPLICATION NUMBER: US 746,121
FILING DATE: 15-AUG-1991
PRIOR APPLICATION DATA:
                                                                                                                                                                                                                                                                                FILING DATE: 22-DEC-1989
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 609,157
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             REGISTRATION NUMBER: 31,822
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                TELEPHONE: (510) 814-2972
INFORMATION FOR SEQ ID NO: 12
FILING DATE: 22-AUG-1986 PRIOR APPLICATION DATA:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        : 892 amino acids
amino acid
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    SEQUENCE CHARACTERISTICS:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     58; Conservative
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  TOPOLOGY: linear

MOLECULE TYPE: protein

US-08-458-819-12
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 Luann Cserr
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              PCT-US91-07035-12
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        Query Match
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     Matches
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         qq
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  οy
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       g
```

```
/ Match 4.9%; Score 81.5; DB 5; Length 892; Local Similarity 17.1%; Pred. No. 6.2; es 58; Conservative 57; Mismatches 116; Indels 109; Gaps
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               Case No. 2580
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 JMBER: WO PCT/US90/07641
21-DEC-1990
                                                                                                                                                                                                                               PCT/US91/07035
ADDRESSEE: Cetus corporations 1400 Fifty-third Street
                                                                                                                                              COMPUTER: IBM PC compatible OPERATING SYSTEM: PC-DOS/MS-DOS
                                                                                                                                                                                                                                                                                  PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 590,490
FILING DATE: 28-SEP-1990
PRIOR APPLICATION DATA:
                                                                                                                                                                                                                                                                                                                                                                                                                                APPLICATION NUMBER: US 590,213
FILING DATE: 28-SEP-1990
PRIOR APPLICATION DATA:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            US 523,394
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         UMBER: US 143,441
12-JAN-1988
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 US 063,509
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            JMBER: US 899,241
22-AUG-1986
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      JMBER: US 746,121
15-AUG-1991
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        JMBER: US 585,471
20-SEP-1990
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  APPLICATION NUMBER: US 609,157
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           APPLICATION NUMBER: US 557,517
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      US 455,611
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    NAME: Sias Ph.D, Stacey R. REGISTRATION NUMBER: 32,630
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                TELECOMMUNICATION INFORMATION:
                                                                                                                                                                                      WordPerfect 5.0
                                                                                                                                                                                                                                                                                                                                                                                          FILING DATE: 28-SEP-1990 PRIOR APPLICATION DATA:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            APPLICATION NUMBER: US 52
FILING DATE: 15-MAY-1990
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    17-JUN-1987
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         FILING DATE: 22-DEC-1989 PRIOR APPLICATION DATA:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   02-NOV-1990
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            FILING DATE: 24-JUL-1990
ATTORNEY/AGENT INFORMATION:
                                                                                                                             Floppy disk
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           REFERENCE/DOCKET NUMBER:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     TELEPHONE: 415-420-3300 INFORMATION FOR SEQ ID NO:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         SEQUENCE CHARACTERISTICS:
                                                                                                                                                                                                         CURRENT APPLICATION DATA:
                                                                                                                                                                                                                                              19910930
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  PRIOR APPLICATION DATA: APPLICATION NUMBER: U
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  PRIOR APPLICATION DATA:
APPLICATION NUMBER: C
FILING DATE: 12-JAN-1
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            PRIOR APPLICATION DATA: APPLICATION NUMBER: W
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     FILING DATE: 02-NOV-1 PRIOR APPLICATION DATA:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         PRIOR APPLICATION DATA:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       PRIOR APPLICATION DATA:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 PRIOR APPLICATION DATA:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        MOLECULE TYPE: protein
                                                                                                        COMPUTER READABLE FORM: MEDIUM TYPE: Floppy
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              PRIOR APPLICATION DATA:
                                                                                                                                                                                                                                                                       CLASSIFICATION: 435
                                                                                                                                                                                                                               APPLICATION NUMBER:
FILING DATE: 199109
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           APPLICATION NUMBER:
                                                                                                                                                                                                                                                                                                                                                                     APPLICATION NUMBER:
FILING DATE: 28-SE
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   APPLICATION NUMBER:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            APPLICATION NUMBER:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      APPLICATION NUMBER:
                                                                    California
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                AMINO ACID
                                              Eneryville
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         linear
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           FILING DATE:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      FILING DATE:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    FILING DATE:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 FILING DATE:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                FILING DATE:
                                                                                     94608
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             PCT-US91-07035-12
```

ą	414 LNPDEK		
λλ	91 KEFREI	KEFREIGELVKTHQLRTSFHPNQFTL	
ą	474 RIFRKL	RIFRKLGRRIYENEMEKLFYEIEMPLIDVLSEMELNGVYFDEEYLKELSKKYQEKMDGIK 533	
දු දි	117FTS 534 EKVFEI	FTSPKESVTKNAVTDMAYHYRMLEAMGIADFSVINIHIGGAYGNKOTATAQFHQNIK 173 	
λ G	174 QLPQEI : 579 VLEELS	QLPQEIKEFMTLENDDKTYTTEFLLQVCEQEDVPFVFDFHH 214 :	
ζζ	215FYAN : 634 RLSSSN	FYANPDDHADLNVALPRMIKTWERIGLQPKVHLSSPKSEQAI 256 : : : : : : RLSSSNPNLQNLPTRSEGKEIRRAVRPQRQDWWILGADYSQIELRVLAHVSKDENLL 691	
oy Ob	257 RSHADYVDAN- :: :: : ! : 692 KAFKEDLDIHT	RSHADYVDANFLLERFROWGTNIDFMI 283 :: :: : : : : :	
RESOLLT US-SCIUM Sequence Pate Pate Pate Pate Pate Pate Pate Pat	SULT 15 -07-854-5968-35 -07-854-5968-35 -08-45968-35 -08-400-400-400 -08-400-400 -08-400-400 -08-400-400 -08-400-400 -08-400-400 -08-4	PRESULT 15 15-15 15-26-35 15-26-35 15-26-35 15-26-35 15-26-35 15-26-36 15-26-36 15-26-36 15-26-36 15-26-36 16-26-3	
Qu Be	Query Match Best Local Si Matches 68;	<pre>Match Local Similarity 20.1%; Pred. No. 9.7; Los Similarity 20.1%; Pred. No. 9.7; les 6%; Conservative 35; Mismatches 112; Indels 123; Gaps 13;</pre>	

ò

```
192 YTTEETLQVCEQEDVPFVFDFHHFYA----NPDDHADLNVALPRMIKTWERIGLQPKVH 246
                                                                                                                                                                                                                  302 EKPYDPFDRSHLKLFTIKYVDVNTNFLLKSEQLLTASERNLDFRDLYDPRDKA----- 354
200 LKDTKLLKTLA-----1GDTITSQELLAQAQSILNKT------HPGYTIYERDSSI 244
                                                                                 96 ------GELVKTHQLRTSFHPN-QFTLFTSPKESVTKNAVTD 131
                                           245 V---THDNDIERTILPMDQEFTYHVKNEQAYEINKKSGLNEEINNTDLISEKYYVLKKG 301
                         73 VPLATHPDVMWDFVTPFQKEF-----RE---------
                                                                                                                                                                                                                                                       247 LSSPKSEQAIRSHADYVDANFLLERFROWGTNIDFMIE 284
                                                                                                                                                                                                                                                                          : | : | | | ::
460 ----NSQLVVSVAGTVE-----GTNQDISLK 481
                                                                                                                                                                                                                                                                                                                                      Search completed: January 15, 2002, 13:51:16
Job time: 345 sec
                                                  Db
                                                                                                        QO
                                                                                                                                    Οy
                                                                                                                                                                   qq
                                                                                                                                                                                          Οÿ
                                                                                                                                                                                                                         pp
                                                                                                                                                                                                                                                          Qγ
                                                                                    ΟŊ
                         δ
```

. .

```
January 15, 2002, 13:50:07; Search time 120.01 Seconds (without alignments) 195.661 Million cell updates/sec
                                                                                                                                                                                                                                                               1 MIFRFGFVSNAMSLWDASPA......ELSSIRGVKRIGGGALQWKS 317
                                                                                                                                                                                                                                                                                                                                                                                                                          522463
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                /SIDS8/gcgdata/geneseq/geneseqp/AA1980.
GenCore version 4.5
Copyright (c) 1993 - 2000 Compugen Ltd.
                                                                                                                                                                                                                                                                                                                                                                                                                   Total number of hits satisfying chosen parameters:
                                                                                                                                                                                                                                                                                                                                                                              522463 seqs, 74073290 residues
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries
                                                                                     OM protein - protein search, using sw model
                                                                                                                                                                                                                                                                                                         BLOSUM62
Gapop 10.0 , Gapext 0.5
                                                                                                                                                                                                                                                                                                                                                                                                                                                               Minimum DB seq length: 0
Maximum DB seq length: 2000000000
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         A_Geneseq_1101:*
                                                                                                                                                                                                                      US-09-724-296-37
1662
                                                                                                                                                                                                                    Title:
Perfect score:
                                                                                                                                                                                                                                                                                                           Scoring table:
                                                                                                                                                                                                                                                                  Sequence:
                                                                                                                                                                                                                                                                                                                                                                              Searched:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         Database
                                                                                                                                 Run on:
```

1: /SIDSB/gcddata/geneseqp/AA1981.DAT:*
3: /SIDSB/gcddata/geneseqp/AA1981.DAT:*
4: /SIDSB/gcddata/geneseqp/AA1981.DAT:*
4: /SIDSB/gcddata/geneseqp/AA1982.DAT:*
5: /SIDSB/gcddata/geneseqp/AA1982.DAT:*
5: /SIDSB/gcddata/geneseqp/AA1985.DAT:*
6: /SIDSB/gcddata/geneseqp/AA1985.DAT:*
7: /SIDSB/gcddata/geneseqp/AA1980.DAT:*
8: /SIDSB/gcddata/geneseqp/AA1980.DAT:*
9: /SIDSB/gcddata/geneseq/geneseqp/AA1980.DAT:*
9: /SIDSB/gcgdata/geneseq/geneseqp/AA1980.DAT:*
9: /SIDSB/gcgdata/geneseq/geneseqp/AA1990.DAT:*
20: /SIDSB/gcgdata/geneseq/geneseqp/AA1990.DAT:*
21: /SIDSB/gcgdata/geneseq/geneseqp/AA1990.DAT:*
22: /SIDSB/gcgdata/geneseq/geneseqp/AA1990.DAT:*
22: /SIDSB/gcgdata/geneseq/geneseqp/AA1990.DAT:*
23: /SIDSB/gcgdata/geneseq/geneseqgp/AA1990.DAT:*
24: /SIDSB/gcgdata/geneseq/geneseqgp/AA1990.DAT:*
25: /SIDSB/gcgdata/geneseq/geneseqgp/AA1990.DAT:*
26: /SIDSB/gcgdata/geneseq/geneseqgp/AA1990.DAT:*
27: /SIDSB/gcgdata/geneseq/geneseqgp/AA1990.DAT:*
28: /SIDSB/gcgdata/geneseq/geneseqgp/AA1990.DAT:*
29: /SIDSB/gcgdata/geneseq/geneseqgp/AA1990.DAT:*
20: /SIDSB/gcgdata/geneseq/geneseqg/geneseqgp/AA1990.DAT:*
21: /SIDSB/gcgdata/geneseqg/geneseqgp/AA1990.DAT:*
22: /SIDSB/gcgdata/gen

SUMMARIES

Query Match Length DB ID
5.6 1041

--1de
1de
1de
228
228
228
4del
1ig
NA

Human ORFX ORF2246 Human H37 amino ac Sorting nexin 1. A. thaliana enviro Sequence encoded b Helicobacter bolor	H. pylori ORF hyse Arabidopsis thalia Arabidopsis thalia S. epidermidis ope Helicobacter pylor Profesin involved i	ν τ	dermidis op dermidis op ori outer ma evisiae ess pest resis dermidis op	Emb ila ila ila ila
AAB4248 AAB0375 AAW7709 AAY7797 AAW7302	AAW5533 AAG1449 AAG1449 AAG8157 AAW7303	AAG3923 AAG3923 AAG3923 AAG3671 AAG3671	AAG822 AAG822 AAW172 AAY172 AAW559 AAG815 AAG815	AAWA457 AAB8846 AAU0099 AAU0098 AAU0096 AAU0096
21 21 19 21 21 9	221182	2222222	200	00000000000000000000000000000000000000
7 7 7 7 3 3 1	1004044	. U U O O H 4 0	1255889888888888888888888888888888888888	80 80 61 61 70 70
	, a a a	. თ თ თ თ თ თ თ	884.5 884.5 884.5 884.5 883.5 883.5	• cc cc cc cc cc cc cc
12 13 14 15 17	118 220 222 233	25 25 27 28 29 30	322 332 332 332 332 332 332 332 332 332	0 0 0 4 4 4 4 4 4 5 5 5 5 5 5 5 5 5 5 5

ALIGNMENTS

```
Delta228-UVDE; ultraviolet damage endonuclease; GST signal peptide;
                                                                                                            glutathione-S-transferase signal peptide; uvel+ gene product; UV Irradiation; DNA damage; UV Irradiation damage; photoproduct; absaic site; aplatinum diaduct; mismatched nucleotide pairing; nucleotide alkylation; skin cancer.
                                                                            S. pombe delta228-UV damage endonuclease.
          AAY44499 standard; Protein; 371 AA
                                                                                                                                                                                                                                                                                                                     Avery AM;
                                                                                                                                                                                                                                       99WO-US12910.
                                                                                                                                                                                                                                                              98US-0088521.
                                                                                                                                                                                                                                                                        99US-0134752
                                                      27-MAR-2000 (first entry)
                                                                                                                                                                     Schizosaccharomyces pombe.
                                                                                                                                                                                                                                                                                                                    Doetsch PW, Kaur B,
                                                                                                                                                                                                                                                                                                                                          WPI; 2000-116417/10.
N-PSDB; AAZ29859.
                                                                                                                                                                                                                                                                                             (UYEM-) UNIV EMORY.
                                                                                                                                                                                                                                      08-JUN-1999;
                                                                                                                                                                                          WO9963828-A1
                                                                                                                                                                                                                                                              08-JUN-1998;
                                                                                                                                                                                                                                                                        18-MAY-1999;
                                                                                                                                                                                                                 16-DEC-1999.
                                 AAY44499;
AAY44499
```

```
RESULT
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      qq
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            qq
q
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           ŏλ
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   q
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            οy
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    Öλ
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        a 288 amino acid deletion of the N-terminal of the S. pombe uvel+ gene product. This is expressed in frame with a GST leader sequence to generate a fusion protein. This provides stable endonuclease fragments for cleaving a double-stranded DNA molecule that has a distorted structure resulting from UV radiation damage, a photoproduct, an abasic site, mismatched nucleotide pairing, aplathium diaduct, an intercalated molecule or alkylation of a nucleotide. Uvelp can be used in compositions for internal or topical application and as a therapeutic agent for skin
                                                             The present sequence is delta228-UV damage endonuclease. Delta228-UVDE is
                                                                                                                                                                                                                                                                                                                                                                          A new truncated ultraviolet damage endonuclease for treatment of skin
                                                                                                                                                                                                                                                                                                                                                                                                                                                                   206 VPFVFDFHHFYANP----DDHADLNVALPRMIKTWERIGLQPKVHLSSPKSEQAI---- 256
                                                                                                                                                                                                                                                                                                                               48 tigrdglesvkglgtgnvldliklvewnhnfgihfmrvssdlfpfash--akygytlefa 105
                                                                                                                                                                                                                                                                                                                                                         90 QKEFREIGELVKTHQLRTSFHPNQFTLFTSPKESVTKNAVTDMAYHYRMLEAMGIADR-- 147
                                                                                                                                                                                                                                                                                                                                                                                                                --SVINIHIGGAYGNKDTATAQFHQNIKQLPQEIKERMTLENDDKTYTTEETLQVCEQED 205
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           226 iplvldwhhnivpgtlregsldlmpliptiretwtrkgitqkqhysesadptaisgmkr 285
                                                                                                                                                                                                                                                                               25; Gaps
                                                                                                                                                                                                                                                                                                        34 TERKEALLTVTKANLRNTM----RTLHYIIGHGIPLYRFSSSIVPLATHPDVMWDFVTPF- 89
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     GST signal peptide; glutathione-S-transferase signal peptide; ultraviolet damage endonuclease; UVDE; delta228-UVDE; repair-deficient E. coli strain; UV irradiation; DNA damage; UV radiation damage; photoproduct; abasic site; aplatinum diaduct; mismatched nucleotide pairing; nucleotide alkylation;
                                                                                                                                                                                                                                                                                                                                                                                                                             21.6%; Score 359.5; DB 21; Length 371; 29.7%; Pred. No. 2.6e-28; tive 60; Mismatches 111; Indels 25;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             GST signal peptide and delta228 S. pombe UVDE fusion protein.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    257 RSHADYVDANFLLERFRQWGTNIDFMIEAKQKDKALLRL 295
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                'note= "Encoded by aaaa"
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          'note= "Encoded by gtt"
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  /note= "Encoded by gtt"
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        Location/Qualifiers
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              AAY44500 standard; Protein; 600 AA
                                    Claim 13; Page 53; 133pp; English.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      fusion protein; skin cancer
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  27-MAR-2000 (first entry)
                                                                                                                                                                                                                                                                              83; Conservative
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                Schizosaccharomyces pombe
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       Misc-difference 147
                                                                                                                                                                                                                                                               Similarity
                                                                                                                                                                                                             371 AA;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   Misc-difference 11
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               Misc-difference 85
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              W09963828-A1
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             Synthetic
                                                                                                                                                                                                              Sequence
                                                                                                                                                                                                                                                    Query Match
                                                                                                                                                                                                                                                                  Local
                                                                                                                                                                                     cancers
                                                                                                                                                                                                                                                                               Matches
                                                                                                                                                                                                                                                                                                                                                                                                               148
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   AAY44500
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                δλ
                                                                                                                                                                                                                                                                                                                                                                                   QQ
                                                                                                                                                                                                                                                                                                                                                                                                               ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                  δŏ
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    ò
```

```
The present sequence is a fusion protein comprising the GST signal peptide and delta228 fruncated ultraviolet damage endonuclase (UVDE) from S. pombe. Delta228-UVDE is a 288 amino acid deletion of the N-terminal of the S. pombe uvel+ gene product. This is expressed in frame with a GST leader sequence. Stable endonuclase fragments can be produced for cleaving a double-stranded DNA molecule that has a distorted structure resulting from UV radiation damage, a photoproduct, an abasic site, mismatched nucleotide pairing, aplathum diaduct, an intercalated molecule or alkylation of a nucleotide. Uvelp can be used in compositions for internal or topical application and as a therapeutic agent for skin
                                                                                                                                                                                                                                                                                                                                                                                                                            A new truncated ultraviolet damage endonuclease for treatment of skin
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    206 VPFVFDFHHFYANP----DDHADLNVALPRMIKTWERIGLQPKVHLSSPKSEQAI---- 256
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    90 QKEFREIGELVKTHQLRTSFHPNQFTLFTSPKESVTKNAVTDMAYHYRMLEAMGIADR-- 147
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      148 --SVINIHIGGAYGNKDTATAQFHQNIKQLPQEIKERMTLENDDKTYTTEETLQVCEQED 205
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    34 TERKEALLTVTKANLRNTM----RTLHYIIGHGIPLYRFSSSIVPLATHPDVMWDFVTPF- 89
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                GST signal peptide; ultraviolet damage endonuclease; UVDE; Uvelp; UV irradiation; DNA damage; UV radiation damage; fusion protein; skin cancer; glutathione-S-transferase.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              DB 21; Length 600;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           111; Indels
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              GST signal peptide and S. pombe UVDE fusion protein.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       257 RSHADYVDANFLLERFRQWGTN.IDFMIEAKQKDKALLRL 295
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       21.6%; Score 359.5; DB 2. 29.7%; Pred. No. 5.5e-28;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        Mismatches
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  Disclosure; Page 56; 133pp; English.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  AAY44498 standard; Protein; 828 AA
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           .09
                                                                                                                                                                                                                                                                             Avery AM,
                                                            99WO-US12910.
                                                                                                                                                     99US-0134752.
                                                                                                                        98US-0088521,
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        27-MAR-2000 (first entry)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           Conservative
                                                                                                                                                                                                                                                                          Kaur B,
                                                                                                                                                                                                                                                                                                                                   WPI; 2000-116417/10.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           Local Similarity
                                                                                                                                                                                                           (UYEM-) UNIV EMORY.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         600 AA;
                                                                                                                                                                                                                                                                                                                                                                   N-PSDB; AAZ29860
                                                            08-JUN-1999;
                                                                                                                                                        18-MAY-1999;
                                                                                                                        08-JUN-1998;
                                                                                                                                                                                                                                                                          Doetsch PW,
16-DEC-1999
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           83;
                                                                                                                                                                                                                                                                                                                                                                                                                                                               cancers -
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               Seguence
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  Query Match
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           Matches
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      AAY44498
```

```
7;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 The present sequence is a fusion protein comprising the GST signal peptide and full length ultraviolet damage endonuclease (UVDE) from Schizosaccharomyces pombe. UVDE genes are isolated by introducing a foreign cDNA library into a repair-deficient E. coli strain and selecting for complemented cells by UV irradiation of the transformants. This provides stable endonuclease fragments for cleaving a double-stranded DNA molecule that has a distorted structure resulting from UV pairing, aplating danage, a photoproduct, an abasic site, mismatched nucleotide pairing, aplatinum diaduct, an intercalated molecule or alkylation of a nucleotide. Uvelp can be used in compositions for internal or topical application and as a therapeutic agent for skin cancers.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             new truncated ultraviolet damage endonuclease for treatment of skin
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           90 QKEFREIGELVKTHOLRTSFHPNOFTLFTSPKESVTKNAVTDMAYHYRMLEAMGIADR-- 147
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           ---SVINIHIGGAYGNKDTATAQFHQNIKQLPQEIKERMTLENDDKTYTTEETLQVCEQED 205
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          VPFVFDFHHFYANP----DDHADLNVALPRMIKTWERIGLQPKVHLSSPKSEQAI---- 256
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           683 iplvldwhhnivpgtlregsldlmpliptiretwtrkgitqkqhysesadptaisgmkr 742
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           Gaps
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          34 TERKEALLIVIKANLRNTM---RTLHYIIGHGIPLYRFSSSIVPLATHPDVMWDFVTPF-
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    25;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        Length 828;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          21.6%; Score 359.5; DB 21; Length 29.7%; Pred. No. 8.9e-28; ive 60; Mismatches 111; Indels
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              RSHADYVDANFLLERFRQWGTNIDFMIEAKQKDKALLRL 295
                                                                                                                                                          1..229
/label-_GST_signal_peptide
                                                                                                                                                                                                                                                                                                                    'note= "Encoded by aaaa"
                                                                                                                                                                                                                                                                                                                                                                                    'note" "Encoded by gtt"
                                                                                                                                                                                                                                                                                                                                                                                                                                                /note= "Encoded by gtt"
                                                                                                                                                                                                                   230..828
/note= "S. pombe UVDE"
                                                                                                                               Location/Qualifiers
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           Claim 7; Page 51; 133pp; English
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             Avery AM;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            99WO-US12910
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       99US-0134752
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         98US-0088521
                          Schizosaccharomyces pombe
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  Best Local Similarity 29.7
Matches 83; Conservative
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      Kaur B,
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               2000-116417/10.
                                                                                                                                                                                                                                                                                                                                               Misc-difference 85
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   (UYEM-) UNIV EMORY.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                828 AA;
                                                                                                                                                                                                                                                                                   Misc-difference
                                                                                                                                                                                                                                                                                                                                                                                                                  Misc-difference
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 N-PSDB; AAZ29858
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             W09963828-A1
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         08-JUN-1999;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     08-JUN-1998;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       18-MAY-1999;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             16-DEC-1999
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      Doetsch PW,
                                                               Synthetic
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      cancers -
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                Sequence
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             Query Match
                                                                                                                                                          Peptide
                                                                                                                                                                                                                          Protein
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              257
SO SC CC CC CC CC X P X X P X X P X X P X X P X X P X X P X X P X X P X X P X X P X X P X X P X X P X X P X X P X X P X X P X X P X X P X X P X X P X X P X X P X X P X X P X X P X X P X X P X X P X X P X X P X X P X X P X P X P X P X P X P X P X P X P X P X P X P X P X P X P X P X P X P X P X P X P X P X P X P X P X P X P X P X P X P X P X P X P X P X P X P X P X P X P X P X P X P X P X P X P X P X P X P X P X P X P X P X P X P X P X P X P X P X P X P X P X P X P X P X P X P X P X P X P X P X P X P X P X P X P X P X P X P X P X P X P X P X P X P X P X P X P X P X P X P X P X P X P X P X P X P X P X P X P X P X P X P X P X P X P X P X P X P X P X P X P X P X P X P X P X P X P X P X P X P X P X P X P X P X P X P X P X P X P X P X P X P X P X P X P X P X P X P X P X P X P X P X P X P X P X P X P X P X P X P X P X P X P X P X P X P X P X P X P X P X P X P X P X P X P X P X P X P X P X P X P X P X P X P X P X P X P X P X P X P X P X P X P X P X P X P X P X P X P X P X P X P X P X P X P X P X P X P X P X P X P X P X P X P X P X P X P X P X P X P X P X P X P X P X P X P X P X P X P X P X P X P X P X P X P X P X P X P X P X P X P X P X P X P X P X P X P X P X P X P X P X P X P X P X P X P X P X P X P X P X P X P X P X P X P X P X P X P X P X P X P X P X P X P X P X P X P X P X P X P X P X P X P X P X P X P X P X P X P X P X P X P X P X P X P X P X P X P X P X P X P X P X P X P X P X P X P X P X P X P X P X P X P X P X P X P X P X P X P X P X P X P X P X P X P X P X P X P X P X P X P X P X P X P X P X P X P X P X P X P X P X P X P X P X P X P X P X P X P X P X P X P X P X P X P X P X P X P X P X P X P X P X P X P X P X P X P X P X P X P X P X P X P X P X P X P X P X P X P X P X P X P X P X P X P X P X P X P X P X P X P X P X P X P X P X P X P X P X P X P X P X P X P X P X P X P X P X P X P X P X P X P X P X P X P X P X P X P X P X P X P X P X P X P X P X P X P X P X P X P X P X P X P X P X P X P X P X P X P X P X P X P X P X P X P X P X P X P X P X P X P X P X P X P X P X P X P X P X P X P X P X P X P X P X P X P X P X
```

```
enzyme
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  The present sequence shows a DNA repair enzyme which specifically recognises a cyclobutane-type dimer and (6-4)-bound product induced, respectively from TT and TC sequences in DNA, by exposure to UV light. The enzyme cleaves these sequences at the 5'-side. The DNA repair enzyme can be used in, e.g. suntan cream for prevention and treatment of skin cancer. Conventional endonuclease acts on either the cyclobutane-type dimer or the (6-4)-bound product in damaged DNA, while the enzyme of the invention can act on both of them.
                                                                                                                                                                                                                          DNA repair enzyme; cyclobutane-type dimer; (6-4)-bound product; exposure; UV Light; ultraviolet; suntan cream; prevention; treatment;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    | : || || || || : || :|| || :| :|| || :|| 336 hrltthpgqfftqlgsprkevvesairdleyhdellsllklpeqqnrdavmiihmggqfgd 395
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   46 ANLRNTMRTLHYIIGHGIPLYRFSSSIVPLATHPDVMWDFVTPFQKE-FREIGELVKTHQ 104
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  105 LRTSFHPNQFTLFTSPKESVTKNAVTDMAYHYRMLEAMGIADR----SVINIHIGGAYGN 160
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               161 KDTATAQFHQNIKQLPQEIKERMTLENDDKTYTTEETLQVCEQEDVPFVFDFHHFYANPD 220
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  D----HADLNVALP----RMIKTWERIGLQPKVHLSSPKSEQAI-----RSHADYVDANF 267
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      26; Gaps
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       DNA repair enzyme and related DNA - specifically recognises cyclobutane-type dimer and (6-4)-bound product induced by exposure
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       Length 656;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        Indels
|:|:| | | :| | | | :| ||||||||||| | 713 rahsdrv-----idfppcdptmdlmieakekeqavfel 775
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   51; Mismatches 105;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  20.8%; Score 345; DB 17; 32.3%; Pred. No. 1.9e-26;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      Disclosure; Page 10-13; 18pp; Japanese.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 LLERFROWGTNIDFMIEAKOKDKALLRLM 296
                                                                                                                                                                                             Neurospora crassa DNA repair enzyme.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  AAW01618 standard; Protein; 656
                                                                                                                                                                                                                                                                                                                                                                                          95JP-0094137.
                                                                                                                                                                                                                                                                                                                                                                                                                         95JP-0094137.
                                                                                                                                                              (first entry)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   87; Conservative
                                                                                                                                                                                                                                                                                                                                                                                                                                                         (SHIS ) SHISEIDO CO LTD.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      WPI; 1996-512663/51
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   Local Similarity
                                                                                                                                                                                                                                                                                            Neurospora crassa.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    656 AA;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       N-PSDB; AAT58286
                                                                                                                                                                                                                                                                                                                       JP08266275-A
                                                                                                                                                                                                                                                                                                                                                                                        29-MAR-1995;
                                                                                                                                                                                                                                                                                                                                                                                                                         29-MAR-1995;
                                                                                                                                                           20-AUG-1997
                                                                                                                                                                                                                                                            skin cancer.
                                                                                                                                                                                                                                                                                                                                                        15-0CT-1996
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        to UV light
                                                                                                                              AAW01618;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    Seguence
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    Query Match
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   Matches
                                                                            AAW01618
                                                             RESULT
                                                                                                                            qq
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               qq
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 δλ
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                Q
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                ò
```

ò g ò q ογ q

qq ò

g

441 pahlregtldisdpklgeriantwkrkgikgkmhysep-cdgavtprdrrkhrprv--- 495

```
The present sequence is delta228-UV damage endonuclease from N. crassa. Delta228-UVDE is a 288 amino acid deletion of the N-terminal of the Uvel+ gene product. This is expressed in frame with a GST leader sequence to generate a fusion protein. This provides stable endonuclease fragments for cleaving a double-stranded DNA molecule that has a distorted structure resulting from UV radiation damage, a photoproduct, an abasic site, mismatched nucleotide pairing, aplatinum diaduct, an intercalated molecule or alkylation of a nucleotide. Uvelp can be used in compositions for internal or topical application and as a therapeutic agent for skin
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      321 hrltthpgqftqlgsprkevvesairdleyhdellsllklpeqqnrdavmiihmgqqfgd 380
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      161 KDTATAQFHQNIKQLPQEIKERMTLENDDKTYTTEFTLQVCEQEDVPFVFDFHHFYANPD 220
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 105 LRTSFHPNQFTLFTSPKESVTKNAVTDMAYHYRMLEAMGIADR----SVINIHIGGAYGN 160
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                221 D----HADLAVALP----RMIKTWERIGLOPKVHLSSPKSEQAI-----RSHADYVDANF 267
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         new truncated ultraviolet damage endonuclease for treatment of skin
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    46 ANLRNTMRTLHYIIGHGIPLYRFSSSIVPLATHPDVMWDFVTPFQKE-FREIGELVKTHQ 104
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   277 anardivkmlcwf-----pfashp-vhgyklapfasevlaeagrvaaelg 320
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        Gaps
                                                                                                                                                                                                                                                                                                                                                                                                                               Delta228-UVDE; ultraviolet damage endonuclease; GST signal peptide;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     41;
                                                                                                                                                                                                                                                                                                                                                                                                                                                           glutathione-S-transferase signal peptide; uvel+ gene product; UV irradiation; DNA damage; UV radiation damage; photoproduct; abasic site; aplatinum diaduct; mismatched nucleotide pairing;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      DB 21; Length 626;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          18.9%; Score 313.5; DB 21; Lengua
30.5%; Pred. No. 3.1e-23;
viematches 97; Indels
                                                                                                                                                                                                                                                                                                                                                                    crassa delta228-UV damage endonuclease.
511 --mtlppcppdmdlmieakdkegavfelm 537
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             nucleotide alkylation; skin cancer.
                                                                                                                                                                    AAY44502 standard; Protein; 626 AA
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               Claim 16; Page 59; 133pp; English.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      Avery AM;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   99WO-US12910.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 99US-0134752
                                                                                                                                                                                                                                                                                                     (first entry)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        82; Conservative
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   Kaur B,
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              WPI; 2000-116417/10.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         Query Match
Best Local Similarity
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                (UYEM-) UNIV EMORY.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        626 AA;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      Neurospora crassa.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     WO9963828-A1
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   08-JUN-1999;
                                                                                                                                                                                                                                                                                                     27-MAR-2000
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              18-MAY-1999;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   Doetsch PW,
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 16-DEC-1999
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               Sequence
                                                                                                                                                                                                                                   AAY44502;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            Matches
                                                                                                                                        AAY44502
                                                                                                     RESULT
```

```
The present sequence is delta228-UV damage endonuclease from B. subtilis. Delta228-UVDE is a 288 amino acid deletion of the N-terminal of the uval+ gene product. This is expressed in frame with a GST leader sequence to generate a fusion protein. This provides stable endonuclease fragments for cleaving a double-stranded DNA molecule that has a distorted structure resulting from UV radiation damage, a photoproduct, an abasic site, mismatched nucleotide pairing, aplatinum diaduct, an intercalated molecule or alkylation of a nucleotide. Uvelp can be used in compositions for internal or topical application and as a therapeutic agent for skin
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             105 LRTSFHPNQFTLFTSPKESVTKNAVTDMAYHYRMLEAMGIADR----SVINIHIGGAYGN 160
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             161 KDTATAQFHQNIKQLPQE1KERMTLENDDKTYTTEETLQVCEQEDVPFVFDFHHFYANPD 220
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  A new truncated ultraviolet damage endonuclease for treatment of skin
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     46 ANLENTMRTLHYIIGHGIPLYRFSSSIVPLATHPDVMWDFVTPFQKE-FREIGELVKTHQ 104
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         277 anardivkmlcwf------pfashp-vhgyklapfasevlaeagrvaaelg 320
                                                                                                                                                                                                                                                             Delta228-UVDE; ultraviolet damage endonuclease; GST signal peptide;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      41;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  18.9%; Score 313.5; DB 21; Length 626; 30.5%; Pred. No. 3.1e-23; ive 49; Mismatches 97; Indels 41;
                                                                                                                                                                                                                                                                          glutathione-S-transferase signal peptide; uvel+ gene product; UV irradiation; DNA damage; UV radiation damage; photoproduct; abasic site; aplatinum diaduct; mismatched nucleotide pairing;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  B. subtillis delta228-UV damage endonuclease.
268 LLERFRQWGTNIDFMIEAKQKDKALLRLM 296
                  --mtlppcppdmdlmleakdkegavfelm 522
                                                                                                                                                                                                                                                                                                                                   nucleotide alkylation; skin cancer.
                                                                                                                       AAY44503 standard; Protein; 626 AA
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        Claim 16; Page 59; 133pp; English.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    Avery AM;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                          99WO-US12910.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             98US-0088521
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            99US-0134752
                                                                                                                                                                                            (first entry)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         Conservative
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    Kaur B,
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  WPI; 2000-116417/10.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     Best Local Similarity
Matches 82; Conserv
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 (UYEM-) UNIV EMORY.
                                                                                                                                                                                                                                                                                                                                                                       Bacillus subtilis.
                                                                                                                                                                                                                                                                                                                                                                                                        W09963828-A1.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 18-MAY-1999;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                           08-JUN-1999;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             08-JUN-1998;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   Doetsch PW,
                                                                                                                                                                                             27-MAR-2000
                                                                                                                                                                                                                                                                                                                                                                                                                                          16-DEC-1999
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     Sequence
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        Query Match
                                                                                                                                                          AAY44503;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      cancers,
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          cancers
                                   496
                                                                                                        AAY44503
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 qq
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   òγ
                                                                                                                                                              δλ
                                   qd
 òχ
```

6

--::

δ

ò q ò g

73 lagddtgaavlthLapqlleaghaftdagvrllmhpeqfivlnsdrpevressvramsah 132

```
D. radiodurans. Delta228-UVDE is a 288 amino acid deletion of the N-terminal of the uvel+ gene product. This is expressed in frame with a GST leader sequence to generate a fusion protein. This provides stable endonuclease fragments for cleaving a double-stranded DNA molecule that has a distorted structure resulting from UV radiation damage, a photoproduct, an abasic site, mismatched nuclectide pairing, applatinum diaduct, an intercalate molecule or alkylation of a nuclectide. Uvelp can be used in compositions for internal or topical application and as a
381 kaatlerfkrnyarlsgscknrlvlenddvgwtvhdllpvceelnipmvldyhhhnicfd 440
                                                                   221 D----HADLNVALP----RMIKTWERIGLOPKVHLSSPKSEQAI-----RSHADYVDANF 267
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            new truncated ultraviolet damage endonuclease for treatment of skin
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 Delta228-UVDE; ultraviolet damage endonuclease; GST signal peptide; glutathione-S-transferase signal peptide; uvel+ gene product; UV irradiation; DNA damage; UV radiation damage; photoproduct; abasic site; aplatinum diaduct; mismatched nucleotide pairing; nucleotide alkylation; skin cancer.
                                                                                                                 The present sequence is delta228-UV damage endonuclease from
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 D. radiodurans delta228-UV damage endonuclease.
                                                                                                                                                                                                                                                 496 --mtlppcppdmdlmieakdkegavfelm 522
                                                                                                                                                                                                      268 LLERFRQWGTNIDFMIEAKQKDKALLRLM 296
                                                                                                                                                                                                                                                                                                                                                                                                                                                    Ā
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 therapeutic agent for skin cancers.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                Claim 16; Page 60; 133pp; English.
                                                                                                                                                                                                                                                                                                                                                                                                                                          AAY44505 standard; Protein; 294
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   Avery AM;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     99WO-US12910
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         98US-0088521
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         99US-0134752
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 (first entry)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           Dienococcus radiodurans
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                Kaur B,
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            WPI; 2000-116417/10
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         (UYEM-) UNIV EMORY
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  08-JUN-1999;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         08-JUN-1998;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 27-MAR-2000
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     18-MAY-1999;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            Doetsch PW,
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  16-DEC-1999
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 AAY44505;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    Sequence
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   cancers
```

```
isomaltoligosaccharides based on cycloisomaltoheptose; (2) substrate specificity; acts on dextran having 1,6 alpha-linkages as main chains, but not on amylopectin, pullulan; and (3) optimum pH and stable pH range; active around pH 6.0 and stable in the range of pH 4.5 to 9.0. The DNA sequence may be used to produce cyclic-isomaltoligosaccharide synthase in microorganisms of the genus Escherichia, particularly E.coli. The resultant cyclic-isomaltoligosaccharide sources.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           physicochemical properties: (1) action: acts on glucose polymers (e.g. dextran) consisting of 1,6-alpha-linkages to form cyclic-
136 YRMLEAMGIADRSVINIHIGGAYGNKDTATAQFHQNIKQLPQEIKERMTLENDDKTYTTE 195
                               133 arvmdglgla-rtpwnll1--lhggkggrgaelaalipdlpdpvrlrlglenderayspa 189
                                                                                          The present sequence represents a novel cyclic-isomaltoligosaccharide synthase isolated from a Bacillus sp. (possibly Bacillus circulans U-155). The cyclic-isomaltoligosaccharide synthase has the following
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  Gaps
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               95 EIGELVKTHQLRTSFHPNQFTLFTSPKESVTKNAVTDMAYHYRMLEAMG--IADRSVIN- 151
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   Cyclic-iso:malt:oligosaccharide synthase - used in the production of
                                                                    196 ETLQVCEQEDVPFVFDFHHFYANP--DDHADLNVA--LPRMIKTWERIGLQPK----VHL
                                                                                                                                                                248 SS----PKSEQAIRSHADYVDANFLLERFRQWGTNIDFMIEAKQKDKAL--LRLM 296
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                99
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  efficiently produce cycloisomaltoheptose which has very strong
                                                                                                                                                                                                                                                                                                                                                                                                       Cyclic-isomaltoligosaccharide synthase; cycloisomaltoheptose; anti decay activity; Escherichia coli; E.coli; microorganism;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           6.2%; Score 103; DB 18; Length 964;
22.1%; Pred. No. 0.24;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              84; Indels
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    cyclo:iso:malto:heptose with strong anti-decay activity
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              44; Mismatches
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         (NORQ ) NORINSUISANSHO SHOKUHIN SOGO.
(NODA ) ZH NODA SANGYO KAGAKU KENKYUSHO.
                                                                                                                                                                                                                                                                                                                                                                        Cyclic-isomaltoligosaccharide synthase.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     Claim 3; Page 8-10; 12pp; Japanese.
                                                                                                                                                                                                                                                                 AAW32619 standard; Protein; 964 AA.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        96JP-0045103.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         96JP-0045103
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              22.1%;
                                                                                                                                                                                                                                                                                                                                     04-FEB-1998 (first entry)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            55; Conservative
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            1997-497314/46.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       anti-decay activity.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               Best Local Similarity
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          964 AA;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                N-PSDB; AAT92376.
                                                                                                                                                                                                                                                                                                                                                                                                                                            recombinant DNA.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                JP09234073-A
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         01-MAR-1996;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        01-MAR-1996;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   09-SEP-1997.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                             Bacillus sp.
                                                                                                                                                                                                                                                                                                  AAW32619;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          Sequence
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           Query Match
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            Matches
                                                                                                                                                                                                                                               AAW32619
 ò
                                 qq
                                                                     ò
                                                                                                       qq
                                                                                                                                        δý
                                                                                                                                                                         qq
                                                                                                                                                                                                                                                                                                  δ
```

14;

13;

Gaps

39;

119; Indels

DB 21; Length 294;

17.6%; Score 292.5; DB 29.8%; Pred. No. 1.4e-21 49; Mismatches

Conservative

Best Local Similarity Matches 88; Conserv

Query Match

VMWD-----FVTPFQKEFREIGELVKTHQLRTSFHPNQFTLFTSPKESVTKNAVTDMAYH 135

81

g à 21 KTLTFARYSKLSKTERKEALLTVTKANLRNTMRTLHYIIGHGIPLYRFSSSIVPLATHPD 80

```
polypeptide. Thes polypeptides are useful for screening for specific inhibitors or activators, in vaccines to induce an immunological response, to raise antibodies and for treating conditions associated with a lack of iles protein. Iles antagonists are used to treat conditions requiring inhibition of iles, specifically infection by C. trachomatis, including ocular or genital trachoma, lymphogranuloma venerium, keratitis, corneal infiltration and prostatitis etc. Iles polypeptides, polynucleotides and their antaqonists can also be used to prevent adhesion of bacteria, especially Gram-positive, to extracellular
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  Isoleucyl tRNA synthetase; ileS; vaccine; ocular; genital trachoma; lymphogranuloma venerium; keratitis; corneal infiltration; prostatitis; bacterial adhesion; in-dwelling device; cell invasion.
152 -----IHIGGAYGNKDTATAQFHQNIKQLPQ-EIKERMTLENDDKTY-----TTEET- 197
                                                                                      612 dagkqfiniagvntspdgaantvwhmskrtpeynilhlinlvnndqnwrnsgnqptaqtn 671
                                                                                                                          ----RMIKTWE 237
                                                                                                                                                            672 latkvyigaeetitgv----yaaspdhnggatqslpfttgtdssgsyisftvpsleyws 726
                                                                                                                                                                                                238 RIGLQPKVHLSSP-----KSEQAIRS-----HADYVDANFLLERFRQWGTNIDFMIEA 285
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         New Chlamydia trachomatis isoleucyl tRNA synthetase polypeptide and polynucleotide - useful as diagnostic reagents and for prevention and treatment of Chlamydia infections which cause ocular trachoma
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      The present sequence represents an isoleucyl tRNA (ileS) synthetase
                                                                                                                                                                                                                 ----LOVCEOEDVPFVFDFHHFYANPDDHADLNVALP----
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                /note= "encoded by GTG"
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    An isoleucyl tRNA (ileS) synthetase.
                                                                                                                                                                                                                                                                                                                                                                                             AAW97216 standard; Protein; 1040 AA
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              Location/Qualifiers
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      Reichard R;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    Claim 6; Page 9-10; 34pp; English.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               (SMIK ) SMITHKLINE BEECHAM CORP. (SMIK ) SMITHKLINE BEECHAM PLC.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             97US-0898798.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           98EP-0305502.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              Chlamydia trachomatis.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      Lawlor EJ,
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        WPI; 1999-097780/09.
                                                                                                                                                                                                                                                                                          11 :11
784 sskddyvlr 792
                                                                                                                                                                                                                                                                    286 KOKDKALLR 294
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 Misc-difference 1
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           N-PSDB; AAX15503
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           10-JUL-1998;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             23-JUL-1997;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                  06-MAY-1999
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      EP893495-A2
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          27-JAN-1999
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      Brown JR,
                                                                                                                                                                                                                                                                                                                                                                                                                              AAW97216;
                                                                                                                                                                                                                                                                                                                                                                                                                                g
                                                                                        g
                                                                                                                            δ
                                                                                                                                                                                                δý
                                                                                                                                                                                                                                                                      ŏ
```

```
13;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      This represents a isoleucyl tRNA synthetase (ileS) polypeptide derived from Chlamydia trachomatis. The ileS polypeptide can be expressed by standard recombinant methodology. The protein produced may then be used to screen for antiblotics active against C. trachomatis and to study the
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               Nucleic acids encoding Chlamydia trachomatis isoleucyl tRNA synthetase polypeptides useful for treating, for example, urethritis, proctitis
matrix proteins in wounds or on in-dwelling devices, and to block iles polypeptide-mediated cell invasion or normal progression of infection.
                                                                                                                                                                                                   ------QNIKQLPQEIKERMTLENDDKTY---TTEETL 198
                                                                                                                                                                                                                                                                                                                                                                                          199 QVCEQEDVPFVFDFHHFYANPDDHADLNVALPRMIKTWERIGLQPKVHLSSPKSEQAIRS 258
                                                                                                                                                                                                                                                                                                                                                                                                                          395 pfcwrtdspliykavnsw----fvavekvkskmlkanesihwtpe-hikggrfgkwleg 448
                                                                                                                                                                              ---- PFQKEFREIGELVKTHQLRTSFHP 111
                                                                                                                                                                                                                                                 112 NQFTLFTSPKESVTKNAVTDMAYHYRMLEAMGLADRSVIN-----IHIGGAYGNKDTA 164
                                                                                                                                                                                                                                                                                     294 -lfpyfqdkke-----leafrilpadfieesegtgivhmapafgeadff 336
                                                                                                                                                                                                                                                                                                                                                         337 acqehnvplvcpvdnqgcytaevkdfvgeyiksadkgiarr--lknenklfyqgtvrhry 394
                                                                                                                                              Gaps
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               Isoleucyl tRNA synthetase: ileS: antibiotic; antibacterial; vaccine; ocular trachoma; Lymphogranuloma Venerium; inclusion conjunctivitis; mucopurulent rhinitis; urethritis; tropical bubo; cervical follicle; prostatitis; proctitis; infant pneumonitis.
                                                                                                          DB 20; Length 1040;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        259 HADYVDANFLLERFRQWGTNIDFMIEAKQKDKALLRL--MDELSSIRGVK 306
                                                                                                                                              42; Mismatches 104; Indels
                                                                                                       Query Match 5.6%; Score 93.5; DE Best Local Similarity 19.7%; Pred. No. 2.6;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        Claim 1; Columns 9-11; 21pp; English.
                                                                                                                                                                                  71 SIVPLATHPDVMW------DFVT----
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             AAY56782 standard; Protein; 1041 AA.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           Reichard RW;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     C. trachomatis iles polypeptide.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         (SMIK ) SMITHKLINE BEECHAM CORP.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       97US-0898978.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    97US-0898978
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 (first entry)
                                                                                                                                                  Conservative
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              Chlamydia trachomatis.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             Lawlor EJ,
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               WPI; 2000-096135/08.
                                                          1040 AA;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 N-PSDB; AAZ46552.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     and prostatitis
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    23-JUL-1997;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       23-JUL-1997;
                                                                                                                                                                                                                                                                                                                       165 TAQFH---
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 27-MAR-2000
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                JS6001602-A.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  14-DEC-1999.
                                                                                                                                                57;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             Brown JR,
                                                          Sequence
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 AAY56782;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           AAY56782
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                qq
                                                                                                                                                                                                                                                                                                                                                            Пp
                                                                                                                                                                                                                                                                                                                                                                                                                              Op
                                                                                                                                                                                                                                                                                                                                                                                                                                                                     ōλ
       55 × 8
                                                                                                                                                                                                                                                       {}_{\rm Q}{\rm Y}
                                                                                                                                                                                                                                                                                                                           ò
```

```
13;
role of the ileS protein in the pathogenesis of disease. They are particularly useful as antigens in the production of antibodies specific for c. trachomatis. The antibiotics and antibodies obtained are useful for treating a wide range of diseases including classic ocular trachoma, Lymphogranuloma Venerium, inclusion conjunctivitis, mucopurulent procutitis, urethritis, tropical bubo, cervical follicles, prostatitis, procedits and infant pneumonitis. The antibodies may also be used to detect the presence of C. trachomatis and related microbes according to
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        Human H37 proteins with a Cdc7 activity regulatory subunit, for controlling cell replication and cell proliferation, useful in treating cancers and diseases due to abnormal production of stem cells \,
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      H37; human; Cdc7 regulatory subunit; cytostatic; proliferative; cancer; anti-proliferative; replication regulator; stem cell.
                                                                                                                                                                                                                                                                                                                    ...----DFVT-----PFQKEFREIGELVKTHQLRTSFHP 111
                                                                                                                                                                                                                                                                                                                                                      234 snmalavhpeltyvrikdkesgdeyilggeslprwfpdresyewigglsgkslvggsyep 293
                                                                                                                                                                                                                                                                                                                                                                                            112 NQFTLFTSPKESVTKNAVTDMAYHYRMLEAMGIADRSVIN-----IHIGGAYGNKDTA 164
                                                                                                                                                                                                                                                                                                                                                                                                                              ------leafrilpadfieesegtgivhmapafgeadff 336
                                                                                                                                                                                                                                                                                                                                                                                                                                                                ------QNIKQLPQEIKERMTLENDDKTY---TTEETL 198
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     QVCEQEDVPFVFDFHHFYANPDDHADLNVALPRMIKTWERIGLQPKVHLSSPKSEQAIRS 258
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          395 pfcwrtdspliykavnsw----fvavekvkskmlkanesihwtpe-hikggrfgkwleg 448
                                                                                                                                                                                                                                                                                Indels 87; Gaps
                                                                                                                                                                                                                                          Length 1041;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                259 HADYVDANFLLERFRQWGTNIDFMIEAKQKDKALLRL--MDELSSIRGVK 306
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              449 ardwa----isrnrywgtpipiw---rsddgellvigsigelealsggk 490
                                                                                                                                                                                                                                          5.6%; Score 93.5; DB 21;
19.7%; Pred. No. 2.6;
tive 42; Mismatches 104;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             (NISC-) JAPAN SCI & TECHNOLOGY CORP.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         AAB03759 standard; Protein; 234 AA
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    Human H37 amino acid sequence #2.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     99WO-JP06076
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         98JP-0311408
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  04-OCT-2000 (first entry)
                                                                                                                                                                                                                                              Query Match 5.6%
Best Local Similarity 19.7%
Matches 57; Conservative
                                                                                                                                                   standard assay protocols.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     WPI; 2000-365580/31.
                                                                                                                                                                                                                                                                                                                                                                                                                              294 -lfpyfqdkke----
                                                                                                                                                                                                                                                                                                                    71 SIVPLATHPDVMW-
                                                                                                                                                                                         1041 AA;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   Masai H;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      N-PSDB; AAA53484
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               (MASA/) MASAI H.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               WO200026250-A1.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                  165 TAQFH----
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     01-NOV-1999;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         30-OCT-1998;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               Homo sapiens
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 11-MAY-2000.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 Arai K,
                                                                                                                                                                                           Sequence
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              AAB03759
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                (ARAI/)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              pp
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            ū
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      셤
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  g
                                                                                                                                                                                                                                                                                                                                                                                                                            q
                                                                                                                                                                                                                                                                                                                                                                                                                                                                  ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          ò
           2222222222
                                                                                                                                                                                                                                                                                                                      ò
                                                                                                                                                                                                                                                                                                                                                                                            ò
```

```
4
                                             The present sequence represents a human H37 protein sequence. H37 is a protein with a Cdc7 activity regulatory subunit. The invention relates to two H37 protein and nucleotide sequences. H37 proteins exhibit cytostatic, proliferative, anti-proliferative, and cell replication regulatory activities. The proteins, encoded genes and DNA fragments are useful in treating cancers and other diseases resulting from abnormal production of stem cells. Antibodies directed against one of the H37 proteins can be used to inhibit cell proliferation.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      immunostimulant, thrombolytic; coaquiant, vasotropic; antidiabetic; hypotensive; dermatological; immunosuppressive; antinflammatory; antiviral; antibacterial; antifungal; antirheumatic; antithyroid; antianaemic; gene therapy; cancer; proliferative disorder; hypertension; neurodegenerative disorder; osteoarthritis; graft vs host disease; cardiovascular disease; diabetes mellitus; hypothyroidism; SCID; AIDS;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           Human: open reading frame, ORFX, detection; cytostatic; hepatotropic; vulnerary; antipsoriatic; antiparkinsonian; nootropic; neuroprotective; anticonvulsant; osteopathic; antiarthritic; immunosuppressant; cardiant;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                severe combined immunodéficiency; malaria; autoimmune disorder; asthma; allergy; aplastic anaemia; nocturnal haemoglobinuria; burn; wound; bone damage; cartilage damage; antiinflammatory disease; coagulation; thrombosis; contraceptive.
                                                                                                                                                                                                                                                                                                                                                                           227 VALPRMIKTWERIGLQPKVHLSSPKS------EQAIRSHADYVDANFLLERFR 273
                                                                                                                                                                                                                                                                                                                                                                                                                 105 vpspesaytaettsphpshdgssfkspdtvclsrgkllvekaikdh-dfipsnsilsnal 163
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               cholesterol ester storage; systemic lupus erythematosus; infection;
                                                                                                                                                                                                                                                                                                                                      19;
                                                                                                                                                                                                                                                                                                DB 21; Length 234;
                                                                                                                                                                                                                                                                                                                                      Indels
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          Human ORFX ORF2246 polypeptide sequence SEQ ID NO:4492.
                                                                                                                                                                                                                                                                                                                                                                                                                                                           274 QWGTNI----DFMIEAKQKDKALLRLMDELSSIR-GVKRIGGGA 312
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                164 swgvkilhiddiryyiegkkkelyllkksstsvrdggkrvgsga 207
                                                                                                                                                                                                                                                                                                                                      42;
                                                                                                                                                                                                                                                                                              5.6%; Score 92.5; DB
31.7%; Pred. No. 0.33;
Live 10; Mismatches
                Claim 2; Page 46-47; 55pp; Japanese.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     AAB42482 standard; Protein; 674 AA
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             31-MAR-1999; 99US-0127607.
02-APR-1999; 99US-0127636.
05-APR-1999; 99US-0127728.
30-MAR-2000; 2000US-0540763.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           31-MAR-2000; 2000WO-US08621.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  08-FEB-2001 (first entry)
                                                                                                                                                                                                                                                                                                                                          33; Conservative
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           Shimkets RA, Leach M;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     (CURA-) CURAGEN CORP
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 WPI; 2000-602362/57.
                                                                                                                                                                                                                                                                                                                   Best Local Similarity
                                                                                                                                                                                                                                        234 AA;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     N-PSDB; AAC76691
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            WO200058473-A2.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        Homo sapiens
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    05-0CT-2000
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               AAB42482;
                                                                                                                                                                                                                                        Sequence
                                                                                                                                                                                                                                                                                                  Query Match
                                                                                                                                                                                                                                                                                                                                          Matches
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    AAB42482
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 QQ
ò
                                                                                                                                                                                                                                                                                                                                                                                                                       Op
                                                                                                                                                                                                                                                                                                                                                                                                                                                               δλ
```

```
which represent the human ORRX open reading frames 1 to 316. Un ARBY 237, which represent the human ORRX open reading frames 1 to 316. Un ARBY 237, sequences have activities such as: cytostatic; hepatoropic; vulnerary; antipornalizic antipornial antiarhritic; immunosuppressant; immunostimulant; cardiant; thrombolytic; coagulant; vasotropic; antidabetic; hypotensive; dermatological; immunosuppressive; antidabetic; hypotensive; dermatological; antistungal; antisheumatic; antidabetic; hypotensive; dermatological; antisheumatic; antithrong; and antianaemic. The sequences can be used for determining the presence of or predisposition to, or preventing or treating pathological conditions associated with an OREX-associated disorder. The nucleic acids can be used to express OREX proteins in gene therapy vectors. The proteins and nucleic acids may be used to treat cancers, proliferative disorders, neurodegenerative disorders, osteoarthritis, graft vs host disease, cardiovascular disease, diabetes mellitus, sypertension, hypothyroidism, cholesterol ester storage, systemic lupus erythematosus, severe combined immunodeficiency (SCID), AIDS, viral, bullorise and englants and englants and englants and englants and englants and englants and englants and englants and englants and englants and englants and englants and englants and englants and englants and englants and englants and englants and englants and englants and englants and englants and englants and englants and englants and englants and englants and englants and englants and englants and englants and englants and englants and englants and englants and englants and englants and englants and englants and englants and englants and englants and englants and englants and englants and englants and englants and englants and englants and englants and englants and englants and englants and englants and englants and englants and englants and englants and englants and englants and englants and englants and englants and englants and englants and englants and englants and englants
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          H37; human; Cdc7 regulatory subunit; cytostatic; proliferative; cancer; anti-proliferative; replication regulator; stem cell.
                                                                                                                                                                 AAC74446 to AAC77606 encode the proteins given in AAB40237 to AAB43397.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   allergies, aplastic anaemia, burns, wounds, bone and cartilage damage, nocturnal haemoglobinuria, antiinflammatory disease; to enhance coagulation; to inhibit thrombosis; and as a contraceptive.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           227 VALPRMIKTWERIGLQPKVHLSSPKS-------EQAIRSHADYVDANFLLERFR 273
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                105 vpspesaytaettsphpshdgssfkspdtvclsrgkllvekaikdh-dfipsnsilsnal 163
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               Gaps
Novel nucleic acids and peptides derived from open reading frame X, useful for treating e.g. cancers, proliferative disorders, neurodegenerative disorders and cardiovascular disease -
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            19;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  5.6%; Score 92.5; DB 21; Length 674; 31.7%; Pred. No. 1.7;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         42; Indels
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     274 QWGTNI----DFMIEAKQKDKALLRLMDELSSIR-GVKRIGGGA 312
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        Pred. No. 1.7;
10; Mismatches
                                                                                                         Claim 11; Page 3676-3677; 5507pp; English.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           (NISC-) JAPAN SCI & TECHNOLOGY CORP.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 AAB03758 standard; Protein; 674 AA
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     Human H37 amino acid sequence #1.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 99WO-JP06076.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   98JP-0311408
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    04-OCT-2000 (first entry)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      33; Conservative
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         Best Local Similarity
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  674 AA;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  Arai K, Masai H;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                (MASA/) MASAI H.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 WO200026250-A1.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        ARAI K.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               Homo sapiens
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           01-NOV-1999;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   30-OCT-1998;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         11-MAY-2000.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  Sequence
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            AAB03758;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               Query Match
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     (ARAI/)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            RESULT 13
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    Matches
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          AAB03758
     ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          qq
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           δλ
```

```
4
                                                                                                                                                           protein with a Cdc7 activity regulatory subunit. The invention relates to two H7 protein and nucleotide sequences. H37 proteins exhibit cytostatic, proliferative, anti-proliferative, and cell replication regulatory activities. The proteins, encoded genes and DNA fragments are useful in treating cancers and other diseases resulting from abnormal production of stem cells. Antibodies directed against one of the H37 proteins can ke used to inhibit cell proliferation.
                                                                 in treating
                                                                                                                                                                                                                                                                                                                                                                                                      227 VALPRMIKTWERIGLQPKVHLSSPKS------EQAIRSHADYVDANFLLERFR 273
                                                                                                                                                                                                                                                                                                                                                                                                                            The present sequence represents a human H37 protein sequence. H37 is
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      Nucleic acids encoding sorting nexin proteins - used for recombinant production and regulating expression of the protein, especially in treating unregulated cell growth e.g. cancer
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 Human; sorting nexin; intracellular domain; cell surface receptor;
                                              Human H37 proteins with a Cdc7 activity regulatory subunit, for
                                                                                                                                                                                                                                                                                                                                                                         19;
                                                                 controlling cell replication and cell proliferation, useful in
cancers and diseases due to abnormal production of stem cells
                                                                                                                                                                                                                                                                                                                                      5.6%; Score 92.5; DB 21; Length 674; 31.7%; Pred. No. 1.7;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             /note= "Xaa not specified, encoded by TYT"
                                                                                                                                                                                                                                                                                                                                                                         42; Indels
                                                                                                                                                                                                                                                                                                                                                                                                                                                                     274 QWGTNI----DFMTEAKQKDKALLRLMDELSSIR-GVKRJGGGA 312
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          translocation; lysosome; down-regulation; cancer.
                                                                                                                                                                                                                                                                                                                                                                       10; Mismatches
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    Claim 3; Column 23-28; 25pp; English
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                Location/Qualifiers
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   AAW77094 standard; Protein; 522 AA.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           Cadena DL, Gill GN, Kurten RC:
                                                                                                               Claim 1; Fig 5; 55pp; Japanese.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               96US-0625322.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             96US-0625322
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  09-NOV-1998 (first entry)
                                                                                                                                                                                                                                                                                                                                                                       33; Conservative
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            (REGC ) UNIV CALIFORNIA.
WPE; 2000-365580/31.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          WPI; 1998-505652/43.
                                                                                                                                                                                                                                                                                                                                                       Best Local Similarity
                                                                                                                                                                                                                                                                                          674 AA;
                  N-PSDB; AAA53483.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         N-PSDB; AAV48248.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  Sorting nexin 1.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              Modified-site
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             01-APR-1996;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 Homo sapiens
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            01-APR-1996;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              US5804412-A.
                                                                                                                                                                                                                                                                                          Sednence
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 AAW77094;
                                                                                                                                                                                                                                                                                                                                        Query Match
                                                                                                                                                                                                                                                                                                                                                                       Matches
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       AAW77094
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    RESULT
 οy
```

```
an activated cell surface receptor expressed on a cell and direct translocation of the receptor to a lysosome in the cell. The nucleic acids encoding SNX are useful for the down-regulation of cell surface receptors for degradation when they are no longer required to carry out a signal. The nucleic acids can be used to express the protein in a cell where expression is desired e.g. where cancer is caused by up-regulation of receptors. Antisense constructs can be used to inhibit native
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    Environmental stress; plant; transgenic plant; anaerobic; flooding; cold;
dehydration; drought; heat stress; salinity; osmotolerance.
              nexins (SNX) have the ability to bind an intracellular domain of
                                                                                                                                                                                                                                                                                                                                                  194 yeklsekhsqngfivppppeksligmtkvkvgkedsssaeflekrraalerylqriv--- 250
                                                                                                                                                                                                                                                                                                                                                                                     77 THPDVMWDFVTPFQKEFREIGELVKTHQLRTSFHPNQFTLFTSPKESVTKNAV----TDM 132
                                                                                                                                                                                                                                                                                                                                                                                                                      307
                                                                                                                                                                                                                                                                                                                                                                                                                                                        133 AYHYRMLEAMGIADR----SVINIHIGGAYGNKDTA--TAQFHQNIKQLPQEIKERMTL 185
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              186 ENDDKTYTTEETLQVCEQEDVPFVFDFHHFYANPD------DH----ADLNVALPRM 232
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         Gaps
                                                                                                                                                                                                                                                                                                               28 YSKLSKTER------KEALLTVTKANL-RNTMRTLHYIIGHGIPLYRFSSSIVPLA 76
                                                                                                                                                                                                                                                                                                                                                                                                         308 wfeeklqeveceegrlrklhavvetlvn---hrkelalntaqfakslamlg------
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   IKTWERIGLQPKVHLSSPKSEQAIRSHADYVDANFLLERFRQ-----WGTNIDFMIEAK
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             356 ssedntalsralsglaeveek--ieglhgeganndffllaellsdyirllaivraafdgr
                                                                                                                                                                                                                                                                            Indels 107;
                                                                                                                                                                                                                                          Length 522;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    A. thaliana environmental stress tolerance related protein.
                                                                                                                                                                                                                                        5.5%; Score 91.5; DB 19;
19.3%; Pred. No. 1.4;
tive 54; Mismatches 106;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      (VLAA-) VLAAMS INTERUNIVERSITAIR INST BIOTECHNOG
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     287 QKDKALLRLMDELSSIRGVKRIGGGALQWKS 317
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              AA.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            AAY77979 standard; protein; 558
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   99WO-EP05652
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    98EP-0202634
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  (first entry)
                                                                                                                                     of receptors. Antisense c expression of the protein.
                                                                                                                                                                                                                                                            Best Local Similarity 19.3
Matches 64; Conservative
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           Verbruggen N;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          Arabidopsis thaliana
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    414 mktwgr------
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          WPI; 2000-205726/18.
N-PSDB; AAZ98359.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             WO200008187-A2.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    04-AUG-1998;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 04-AUG-1999;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  14-JUN-2000
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             AAY77979:
                                                                                                                                                                                          Seguence
                                                                                                                                                                                                                                           Query Match
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         Lee JH,
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           15
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        451
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              AAY77979
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       ò
×333333333333
                                                                                                                                                                                                                                                                                                                                                  음
                                                                                                                                                                                                                                                                                                                                                                                     ò
                                                                                                                                                                                                                                                                                                                                                                                                                      g
                                                                                                                                                                                                                                                                                                                                                                                                                                                          ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           셤
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   õ
```

```
involved in tolerance to environmental stress in plants. The sequences (AAZ98305-298365) are useful for producing a transgenic plant having enhanced tolerance or resistance to environmental stress conditions such as anaerobic, flooding, cold, delydration, drought, heat stress or salinity. This is useful for producing improved yield, growth, development and productivity under environmental stress conditions, and also provides growth of crops in areas where they cannot grow without the induced osmotolerance. Sequences AAY77925-984 represent polypeptide sequences from A. thaliana that are encoded by the genes involved in
Isolation of polynucleic acids useful for producing transgenic plant by isolating genes involved in tolerance to environmental stress
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   69 SSSIVPLATHPDV-----MWDFVT--PFQKEFREIGELVKTHQLRTSFHPNQFTLFTS 119
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       120 PKESVTKNAVTDMAYHYRMLEAMGIADRSVINIHIGGAYGNKDTATAQFHQNIKQLPQEI 179
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               ---rvmkalgv----llnvkfggssg-edtemkeaderkepepeme 212
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            234 KTWERIGLOPKVHLSSPKSEQAIRSHADYVDANFLLERFROWGTNIDFMIEA-KOKDKAL 292
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    114 srsrvssksnpfydafggkemwekltadpgtrvylegddfyktmk-eigrnpnnlnlymk 172
                                                                                                          to isolation of coding sequences and/or genes
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         180 KERMILENDDKTYTTEETLQVCEQEDVPF-VFDF----HHFYANPDDHADLNVALPRMI
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 213 pmelteeergkkerkekalkekgegnvaykkkdfgravehytkamelddedisyltnr--
                                                                                                                                                                                                                                                                                                                                                                                                                               DB 21; Length 558;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                           Indels
                                                                                                                                                                                                                                                                                                                                                                                                                                                                         91;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                           45; Mismatches
                                                                                                                                                                                                                                                                                                                                                                                                                               Score 91;
Pred. No. 1
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          Search completed: January 15, 2002, 13:50:09 Job time: 298 sec
                                                                Page 278-282; 312pp; English.
                                                                                                                                                                                                                                                                                                                           environmental stress tolerance.
                                                                                                                                                                                                                                                                                                                                                                                                                                 5.5%; 21.8%;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                           Conservative
                                                                                                            The invention relates
                                                                                                                                                                                                                                                                                                                                                                                                                                   Query Match
Best Local Similarity
                                                                                                                                                                                                                                                                                                                                                                   558 AA;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 173 dk-----
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          318
                                                                                                                                                                                                                                                                                                                                                                                                                                                                           53;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        vkm
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                293
                                                                                                                                                                                                                                                                                                                                                                                                                                                                             Matches
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          271
    δλ
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               q
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               Op
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            δ
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    Db
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              ŏ
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   qq
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  ŏ
                                                                                                                                                                                                                                                                                                                                               16;
```

Gaps

.

GenCore version 4.5 Copyright (c) 1993 - 2000 Compugen Ltd.

OM protein - protein search, using sw model

January 15, 2002, 14:07:07; Search time 45.07 Seconds (without alignments) 533.662 Million cell updates/sec Run on:

Title: Perfect score:

US-09-724-296-36 3482 1 MPSRKSKAAALDTPQSESST.....REVKKGKVPEEVEDEGEFDG 656 Sequence:

BLOSUM62 Gapop 10.0 , Gapext 0.5 Scoring table:

100059 segs, 36664827 residues Searched:

100059

Total number of hits satisfying chosen parameters:

Minimum DB seq length: 0 Maximum DB seq length: 200000000

Post-processing: Minimum Match 0% Maximum Match 100% Listing first 45 summaries

SwissProt_39:* Database : Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

	Description	P45864 bacillus su	saccharom		homo	mus m		уошо	sacc	Q9ji55 cricetulus	Pl2839 rattus norv	Q92794 homo sapien	P12036 homo sapien	rattu	Q60974 mus musculu		082709 pisum sativ	homo	P46100 homo sapien	homo	097159 drosophila	Q9y618 h nuclear r	P40563 saccharomyc	P34926 rattus norv	P29402 arabidopsis	P07197 homo sapien	Q39994 helianthus	Q13459 homo sapien	P19246 mus musculu	Q02952 homo sapien		Q39817 glycine max	9532	O84098 chlamvdia t
SUMMARIES	QI	YWJD BACSU	YFC3_YEAST	DIL2_HUMAN	KI67_HUMAN	NFM_MOUSE	CALD_HUMAN	CENE_HUMAN	ZDS1_YEAST	PLE1_CRIGR	NFM_RAT	MOZ_HUMAN	NFH_HUMAN	NEST_RAT	NCR1_MOUSE	DPOZ_HUMAN	CALX_PEA	CRK7_HUMAN	ATRX_HUMAN	MYH9_HUMAN	CHDM_DROME	NCR2_HUMAN	YIS3_YEAST	MAPA_RAT	CAX1_ARATH	NFM_HUMAN	CALX_HELTU	MY9B_HUMAN	NFH_MOUSE	AKAC_HUMAN	CALD_RAT	CALX_SOYBN	TFE3_HUMAN	IF2_CHLTR
	DB	-	Н	Н	Н	_	-	П	П	-	П	Н	-	~	, - 1	, - 1	٦	-	7	_	~		-	, ⊶1		-	Н	-	 4	_	-	-	-	-
	ouery Match Length DB	320	196	747	3256	848	793	2663	915	4473	845	2004	1020	1805	2453	3130	551	1490	2375	1960	1982	2517	619	2774	530	915	540	2158	1087	1781	531	546	743	892
ď	Query Match	10.0	4.2	4.0	3.9	3.9	3.9	3.8	3.8	3.8	3.8	3.7	3.7	3.7		•	3.6	3.6	3.6	3.6	3.6	3.5	3.5	3.5	3.5	3.5	3.5	3.5	3.5	3.5	3.5	3.4	3.4	3.4
	Score	348.5	146	141	137	136.5	135	134	132.5	131.5	131	128.5	128	128	127.5	127	126.5	126.5	126	124.5	124.5	123	122.5	122.5	122	122	121.5	121.5	121	121	120.5	120	120	119
	Result No.		2	3	4	S	9	7	80	σ	10	11	12	13	14	15	16	17	18	19	20	21	22	23	24	25	56	27	28	29	30	31	32	33

14152 homo sapier	075376 homo sapien	27951 streptococo	99uld4 homo sapier	25980 xenopus lae	21675 homo sapier	008638 mus musculu	11434 xenopus lae	04404 sus scrofa	16884 rattus norv	39730 saccharomy	07261 saccharomýc
	0	Ь	o o	. 6.	d.	0	д	Δ.	Ф	М	д
IF3A_HUMAN	NCR1_HUMAN	BAG_STRAG	BRF3_HUMAN	UBF2_XENLA	T2D1_HUMAN	MYHB_MOUSE	41_XENLA	CMGA_PIG	NFH_RAT	IF2P_YEAST	GCR1_YEAST
_	, ,		7		-	_	_	_	_	-	_
1382	2440	1164	1214	701	1872	1972	801	446	831	1002	844
3.4	3.4	3.4	3.4	3.4	3.4	3.4	3.3	3,3	3.3	3.3	3.3
119	119	118.5	118.5	117.5	117.5	117.5	116.5	116	116	116	115.5

ALIGNMENTS

δŏ

g

This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss institute of Bioinformatics and the EMBL outstation the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See http://www.isb-sib.ch/announce/or send an email to license@isb-sib.ch). 326 KSSFAENSDKKQKKKQVGFASSLEI---HEVENLKEENKRQMQSFAVPMYETQESTG--I 380 63 AGRMMGKDANGHC-----LREGKEQEEGVKMAIEGLARMERRLQRATKRQKKQLEEDGI 116 162 -----VDDVVKMEPAATNIIEPEDA-QDAAERG-------AARPPAVNS-S 198 485 EKEPVVGDVIEKEPAVTDIVEREPAVNDIVERKPVVGDIIEKEPTINDIVEKEPEINSKS 544 199 YLPLPWKGRLGYACLNTYLRNAKPPIFSSRTCRM---ASIVDHRHPLQFEDEPEHHLKNK 255 545 EFETPFKKK-----KLKSLQKP-RSSKSMKKKFDPKILENISDDDYDDDDGNKKLL 595 256 PDKSKEPQDELGHKF---VQELGLANARDIVKMLCWNEKYGIRFLRLSSEMFPFASHPVH 312 596 SNKSKNNTDE-QDKFPSKIQEVSRSMAKT-----GATVGSEPVRITNVDY----HALG 643 Query Match
4.2%; Score 146; DB 1; Length 796;
Best Local Similarity 19.2%; Pred. No. 0.12;
Matches 108; Conservative 97; Mismatches 183; Indels 174; Gaps 381 ANKMTSDEFDGDLFAKMLGVQEADEVHEKYKEELINQERLEGEASRSNRRTR----- 432 117 PVPSVVSRF------PTAPYHHKSTNAEEREA--KEPVLKTHSKDVEREAEIG--- 161 313 G-----YKLAPFASEVLAEAGRVAAELGHRLTTHPGQFTQLGSPRKEVVESAIRDLE 364 ------EDFK 677 4 RKSKAAALDTPQSESST-FSSTLDSSAPSPARNLRRSGRNILQPSSEKDRDHEKRSGEEL 62 Murakami Y., Naitou M., Hagiwara H., Shibata T., Ozawa M., Sasanuma S.-I., Sasanuma M., Tsuchiya Y., Soeda E., Yokoyama K., Yamazaki M., Tsushiro H., Eki T.; "Analysis of the nuclectide sequence of chromosome VI from Saccharomyces cerevisiae."; Eukaryota; Fungi; Ascomycota; Saccharomycotina; Saccharomycetes; 01-NoV-1995 (Rel. 32, Created)
01-NOV-1995 (Rel. 32, Last sequence update)
01-NOV-1995 (Rel. 32, Last annotation update)
HYPOTHETICAL 91.4 KDA PROTEIN IN STE2-FRS2 INTERGENIC REGION. Hypothetical protein. SEQUENCE 796 AA; 91421 MW; 7E2EE2CEB4708BC3 CRC64; Saccharomycetales; Saccharomycetaceae; Saccharomyces. NCBL_TaxID=4932; 644 GNLDDMVKAYSLGLYDDDLEEDPGTIVEKL-------268 LLPLLERFROWGTNIDFMIEAKOKDKALLRLM 299 Saccharomyces cerevisiae (Baker's yeast). STRAIN=S288C / AB972; MEDLINE=95400292; PubMed=7670463; Nat. Genet. 10:261-268(1995). EMBL; D50617; BAA09215.1; -. InterPro; IPR001440; TPR. STANDARD; S0001871; YFL023W YFC3_YEAST P43573: RESULT 2 YFC3_YEAST

Op

δy Op δ

ò

δy qq ò

da o	365 YHDELLSLLKLPEQONRDAVMIIHMGGQFGDKAATLERFKRNYARLSQSCKNRLVLENDD 424 :::: :: :: :: :: 678 EYNKQVELLRDEIRDFQLENKPVTMEEEEN 707 425 VGWTVHDLLPVCEELNIPMVLDYHHHICFDPAHLREGTLDISDPKLQERIANTWKRKGI 484
q Q	08 DGNVMNDIIEHEFPESYINDEDEVALHPGRLQE-EVAIEYRLKEATASKWGS 75
Q.Y	485 KQKMHYSEPCDGAVTPRDR 503 : : : 760SSPAAHTEGELEPIDK 775
	LUT 3 - HUMAN DILZ_HUMAN STANDARD; PRT; 747 AA.
AC DT	090LW0; Q9UL00; Q9Y2M1; Q9UFN9; Q9NRA3; 20-AUG-2001 (Rel. 40, Created)
DT DE	(0
	(DIFFERENTIALLY EXPRESSED IN LUNG CELLS 2) (DIL-2) (TARGETING PROTEIN FOR XKLP2) (CROORET PROTEIN) (CROORET PROTEIN) (PROTEIN FLS3S3).
s S S S S	
8 8 8	ae; Homo.
N X	SECULENCE FROM N.A.
XX	T e + OdoV H Odeway
R R P	mainda N., Avino 1., malsuno 1., makenosinta 5., Auwano n., Dokuda 5., "Identification of genes (SPONO and C20orf2) differentially expressed
RT	
X X V	[2] SEQUENCE FROM N.A.
RX RA	er B., Parware
RT	0 gene (C20orfl) ation.";
R.L.	
7 Z	SCUENCE FROM N.A.
RA	
	"Fetal gene preferentially expressed in colorectal cancer."; Submitted (MAR-1999) to the EMBL/GenBank/DDBJ databases.
R R R	Ansorge W., Wirkner U., Mewes HW., Gassenhuber J., Wiemann S.; Submitted (SEP-1999) to the EMBL/GenBank/DDBJ databases.
	13 SEQUENCE OF 42-747 FROM N.A.
	0
	"TPX2, a novel Xenopus MAP involved in spindle pole organization.";
	CHARACTERIZATION.
	MEDLINE=9/35108L; PubMed=920/45/; Heidebrecht H.J., Buck F., Steinmann J., Sprenger R., Wacker H.H.,
	d nuclea
	and M.";
	ING MITOSIS IT IS STRICTLY
5 5	ND WITH THE

```
This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation the Buropean Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See http://www.isb-sib.ch/announce/or send an email to license@isb-sib.ch).
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                36;
TISSUE SPECIFICITY: EXPRESSED IN LUNG CARCINOMA CELL LINES BUT NOT
                                                                              DEVELOPMENTAL STAGE: EXCLUSIVELY EXPRESSED IN PROLIFERATING CELLS FROM THE TRANSITION G1/S UNTIL THE END OF CYTOKINESIS.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        63 AGRMMGKDANGHCLREGKE--QEEGVKMAI-EGLARMERRLQRATKRQKKQLEEDGIP-- 117
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            99 ACSSLEVEAAISRKTPAOPORRSLRLSAOKDLEOKEKHHVKMKAKRCATPVIIDEILPSK 158
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      168 MEPAATNIIEPEDA----QDAAERGAARPPAVNSSYLPLPWKGRLGYACLNTYLRNAKPP 223
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     283 VKMLCWNEKYGIRFLRLSSEMFPFASHPVHGYKLAPFASEVLAEAGRVAAELGHRLTTHP 342
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           400 LERFKRNYARLSQSC------KNRLVLENDDVGWTVHDLLPVCEELNIPMVLDYH 448
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         118 -----VPSVVSRFPTAPYHHKSTNAEEREAKEPVLKTHSKDVERE--AEIGVDDVV--K 167
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       224 IFSSRTCRMASIVDHRHPLQFEDEPEHHLKNKPDKSKEPQD-ELGHKFVQELGLANARDI 282
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          343 GQFTQLGSPRKEVVESAIRDLEYH---DELLSLLKLPEQQNRDAVMIIHMGGQFGDKAAT 399
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              LRKHPSSPARVTKGCTIVKPFNLSQGKKRTFDE-----TVSTYVPLAQQ-----VEDFH 335
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        449 HHNICFDPAHLREGTLDISDPKLQERIANTWKRKGIKQKMHYSEPCDGAVTPRDRRKHRP 508
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   336 KRTP--NRYHLRSKKDDIN-------LLPSKSSVTKICRDPQTPVLQTKHRA 378
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    ----IGQPVKKSVSQVTKSVDFHFRTDERIKQHPKNQEEYKEV---------NFTSE 286
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 RVMTLPPCPPDMDLMIEAKDKEQAVFELMRTF-----KLPGFEKINDMVPYD- 555
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    RAVT---CKSTAELEAEELEKLOOYKFKARELDPRILEGGPILPKKPPVKPPTEPIGFDL 435
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               -----PVKAPKKKGGKRKR- 579
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     436 EIEKRIQERESKKKTEDEHFEFHSRPCPTKILEDVVGVPEKKVLPITVPKSPAFALKNRI 495
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                Gaps
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       3 SRKSKAAALDTPQSESSTFSSTLDSSAPSPARNLRRSGRNILQPSSEKDRDHEKRSGEEL 62
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 2 SQVKSSYSYDAP-SDFINFSS-LDDEGDT-----QNIDSWFEEK----ANL 41
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     : | :||: || ||: ||: || ||: ||: || ||: ||: ||: ||: ||: ||: ||: ||: ||: ||: ||: ||: ||: ||: ||: ||: ||: ||: ||: ||: ||: ||: ||: ||: ||: ||: ||: ||: ||: ||: ||: ||: ||: ||: ||: ||: ||: ||: ||: ||: ||: ||: ||: ||: ||: ||: ||: ||: ||: ||: ||: ||: ||: ||: ||: ||: ||: ||: ||: ||: ||: ||: ||: ||: ||: ||: ||: ||: ||: ||: ||: ||: ||: ||: ||: ||: ||: ||: ||: ||: ||: ||: ||: ||: ||: ||: ||: ||: ||: ||: ||: ||: ||: ||: ||: ||: ||: ||: ||: ||: ||: ||: ||: ||: ||: ||: ||: ||: ||: ||: ||: ||: ||: ||: ||: ||: ||: ||: ||: ||: ||: ||: ||: ||: ||: ||: ||: ||: ||: ||: ||: ||: ||: ||: ||: ||: ||: ||: ||: ||: ||: ||: ||: ||: ||: ||: ||: ||: ||: ||: ||: ||: ||: ||: ||: ||: ||: ||: ||: ||: ||: ||: ||: ||: ||: ||: ||: ||: ||: ||: ||: ||: ||: ||: ||: ||: ||: ||: ||: ||: ||: ||: ||: ||: ||: ||: ||: ||: ||: ||: ||: ||: ||: ||: ||: ||: ||: ||: ||: ||: ||: ||: ||: ||: ||: ||: ||: ||: ||: ||: ||: ||: ||: ||: ||: ||: ||: ||: ||: ||: ||: ||: ||: ||: ||: ||: ||: ||: ||: ||: ||: ||: ||: ||: ||: ||: ||: ||: ||: ||: ||: ||: ||: ||: ||: ||: ||: ||: ||: ||: ||: ||: ||: ||: ||: ||: ||: ||: ||: ||: ||: ||: ||: ||: ||: ||: ||: ||: ||: ||: ||: ||: ||: ||: ||: ||: ||: ||: ||: ||: ||: ||: ||: ||: ||: ||: ||: ||: ||: ||: ||: ||: ||: ||: ||: ||: ||: ||: ||: ||: ||: ||: ||: ||: ||: ||: ||: ||: ||: ||: ||: ||: ||: ||: ||: ||: ||: ||: ||: ||: ||: ||: ||: ||: ||: ||: ||: ||: ||: ||: ||: ||: ||: ||: ||: ||: ||: ||: ||: ||: ||: ||: ||: ||: ||: ||: ||: ||: ||: ||: ||: ||: ||: ||: ||: ||: ||: ||: ||: ||: ||: ||: ||: ||: ||: ||: ||: ||: ||: ||: ||: ||: ||: ||: ||: ||: ||: ||: ||: ||: ||: ||: ||: ||: ||: ||: ||: ||: ||: ||: ||: ||: ||: ||: ||: ||: ||: ||: ||: ||: ||: ||: ||: ||: ||: ||: ||: ||: ||: ||: ||: ||: ||: ||: ||: ||: ||: ||: ||: ||: ||: ||: ||: ||: ||: ||: ||: ||: ||: ||: ||: ||: ||: ||: ||: ||: ||: ||: ||: ||: ||: ||: ||: ||: ||: ||: ||: ||: ||: ||: ||: ||: ||: ||: ||: ||: ||: ||: ||: ||: ||: ||: ||: ||: ||: ||: ||: ||: ||: ||: ||: ||: ||: ||: ||: ||: ||: ||: ||: ||: ||: ||: ||: ||: ||: ||: ||: ||: ||: ||: ||: ||: ||: ||: ||: ||: ||: ||: ||: ||: ||: ||: ||: 
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      -----AKOKFLKSTEEQELEKSMKMQQE-----V
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   226 VEMRKKNEEF------
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         4.0%; Score 141; DB 1; Length 747;
19.6%; Pred. No. 0.21;
Live 98; Mismatches 233; Indels 258;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       EMBL; AFI...
MIM, 605917; -..
MIM, 605917; -..
MOCLEAT protein; Phosphorylation.
CONCLET 182 182 K -> N (IN REF. 3).
CONCLUT 273 273 K -> E (IN REF. 1).

OFFGR MW; E028E0BB50BBCA0F CRC64;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            BAA76931.1; -. CAB55982.1; -.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  EMBL; AB027467; BAA85893.1; -.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         AAF03248.1; -.
                                           IN NORMAL LUNG TISSUES
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                Matches 144; Conservative

    !- PTM: PHOSPHORYLATED.

                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         Query Match
Best Local Similarity
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             AF098158;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            AB024704;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                AL117534;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  202 -----
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        243
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        379
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      287
   d
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  ô
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            q
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       ογ
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      q
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           qq
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        Dp
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         ò
```

```
This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation-the Buropean Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See http://www.isb-sib.ch/announce/
                                496 RMPTKEDEEEEDEPVVIKAQPVPHYGVPFKP-QIPEARTVEICPFS--FDSRDKERQLQKE 552
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             SUBCELLULAR LOCATION: NUCLEAR. PREDOMINANTLY LOCALIZED IN THE GIPHASE IN THE PERINUCLEOLAR REGION, IN THE LATER PHASES IT IS ALSO DETECTED THROUGHOUT THE NUCLEAR INTERIOR, BEING PREDOMINANTLY LOCALIZED IN THE NUCLEAR MATRIX. IN MITOSIS, IT IS PRESENT ON ALL
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        ALTERNATIVE PRODUCTS: 2 ISOFORMS; A LONG FORM (SHOWN HERE) AND A SHORT FORM; ARE PRODUCED BY ALTERNATIVE SPLICING.
DEVELOMENTAL STAGE: EXPRESSION OF THIS ANTIGEN OCCURS.
PREFERENTIALLY DORING LATE G1, S, G2, AND M PHASES OF THE CELL
CYCLE, WHILE IN CELLS IN G0 PHASE THE ANTIGEN CANNOT BE DETECTED.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     "The cell proliferation-associated antigen of antibody Ki-67: a very large, ubiquitous nuclear protein with numerous repeated elements, representing a new Kind of cell cycle-maintaining proteins."; J. Cell Biol. 123:513-522(1993).
--TTDEEAAEPEEVDTAADDVK--DAPEGPKEVPEEERAMGGPYNRVYWPLGCEEWLKPK
                                                                                                                                                                                                                                                                                                                                                                Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Buteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 Schlueter C., Duchrow M., Wohlenberg C., Becker M.H.G., Key G., Flad H.-D., Gerdes J.;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               FHA.
16 X 122 AA APPROXIMATE REPEATS.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     Cell cycle; Antigen; Muclear protein; ATP-binding; Repeat;
Alternative splicing.
DOMAIN 27 76 FHA.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         Submitted (MAR-1997) to the EMBL/GenBank/DDBJ databases.
-!- FUNCTION: THOUGHT TO BE REQUIRED FOR MAINTAINING CELL
PROLIFERNITHON.
                                                                                                                                                                                                                                                                      01-NOV-1995 (Rel. 32, Last sequence update) 20-AUG-2001 (Rel. 40, Last annotation update)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            or send an email to license@isb-sib.ch).
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            SIMILARITY: CONTAINS 1 FHA DOMAIN.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                   MEDLINE=94043435; PubMed=8227122;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         MART; SM00240; FHA; 1.
SMART; SM00240; FHA; 1.
PROSITE; PS50006; FHA_DOMAIN; 1.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             InterPro; IPR000253; FHA_domain.
                                                                                                                                                                                                                                                  01-NOV-1995 (Rel. 32, Created)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  EMBL; X65550; CAA46519.1; -.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  EMBL; X65551; CAA46520.1; -
EMBL; X94762; CAA64388.1; -
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   SEQUENCE OF 1-31 FROM N.A.
                                                                                                                                                                                                                  STANDARD;
                                                                                                                  553 KKIKELOKGEVPK 565
                                                                            636 K -- REVKKGKVPE 646
                                                                                                                                                                                                                                                                                                                                                Homo sapiens (Human)
                                                                                                                                                                                                                                                                                                                                                                                                                                             SEQUENCE FROM N.A.
                                                                                                                                                                                                                                                                                                                                                                                                           NCBI_TaxID=9606;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            CHROMOSOMES
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             1000
                                                                                                                                                                                                                                                                                                          ANTIGEN KI-67
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         MIM; 176741;
                                                                                                                                                                                                          KI67_HUMAN
P46013;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         Gerdes
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 DOMAIN
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 REPEAT
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     REPEAT
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     REPEAT
                                                                                                                                                                                            KI67_HUMAN
                                                                                                                                                                                                                                                                                                                              MKI67
                                                                                                                                                                        RESULT
                                      q
                                                                            ó
```

```
30;
                                                                                                                                                                                                                                                                                                                                                                                                                       2488 KVDMKEEPLAVSKLT-RTSGETTQTHTEPTGDSKSIKAFKESPKQILDPAASVTGSRRQL 2546
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                2547 RTRKEKARALEDLVDFKELFS----APGHTEESMTIDKNTKIPCKSPPPELTDTATSTKR 2602
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           2603 CPKTRPRKEVKEELSAVERLTQTSGOSTHTHKEPASGDEGIKVLKORAKKKPNPVEEEPS 2662
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     2663 RRRPRAPKEKAQ----PLEDLAGFTELSETSGHTQESLTAGKATK---IPCESPPLEVVD 2715
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                2716 TTASTKRHLRTRVQKVQVKEEPSAVKFTQTSG--ETTDADKEPAGEDK-GIKALKESAKQ 2772
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         2773 TPAPAASVTGSRRRPRAPRESAQAIEDLAGFKDPAAGHTEESMTDDKTTKIPCKSSPELE 2832
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  2833 DTATSSKRRPRTRAQKVEVKEELLAVGKLTQTSGETTHTDKEPVGEGKGTKAFKQPAKRN 2892
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         2928 -------PGHTEELANGAADSFTSAPKQTPDSGKPLKISRRVLRAPK-- 2967
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  ------VEPVGDVVSTRDPVKSQSKSNTSLPPLP-----FKRGGGKDGSVTGTKRLRC 3014
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  2428 KEKAEALEDLVGFKELFQTPGHTEESMTDDKITEVSCKSPQPESFKTSRSSKQRLKIPLV 2487
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       --PEQQN 380
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           ELNIPMVLDYHHHNICFDPAHLRE------GTLDISDP-KLQERIANTWKRK 482
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       483 GIKQKMHYSEPCDGAVTPRDRRKHRPRVMT-LPPCPPDMDLMIEAKDKEQAVFELMRTFK 541
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   247 ---EPEHHLKNKPDKSKEPQDELGHKFVQELGLANARDIVKMLCWNEKYGIRFLRLSSEM 303
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              304 FPFASHPVHGYKLAPFASEVLAE----AGRVAAELGH------RLTTHP----GQFT 346
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  381 RDAVMIIHMGGQ----FGDKAATLERFKRNYARLSQSCKNRLVLENDDVGWTVHDLLLPVCE 437
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                542 LPGFEKINDMVPYDRDDENRPAPPVKAPKKKKGGKRKRRTTDEEAAEPEEVDTAADDVKDA 601
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                -----LKTHSKDVEREAEI-GVDDVVKMEPAA----TNII------EPEDAQDAAE 186
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         187 RGAARPPAVNSSYLPLPWKGRLGYACLNTYLRNAKPPIFSSRTCRMASIVDHRHPLQFED 246
                                                                                                                                                                                                                                                                                                                                                                                          EKDRDHEKRSGEELAGRMMGKDANGHCLREG-----KEQEEGVKMAIEGLARME-RRLQR 102
                                                                                                                                                                                                                                                                                                   Gaps
                                                                                                                                                                                                                                                                                                                                5 KSKAAALD-----TP--QSESSTFSSTLDSSAPSPA----RNLRRSGRNILQPSS 48
                                                                                                                                                                                                                                                                                                   92; Mismatches 296; Indels 200;
                                                                                                                                                                                                                                                                     DB 1; Length 3256;
                                                                                                                                                                                                            SING (IN SHORT ISOFORM).
578F8C51BED42517 CRC64;
                                                                                                                                                                                                                                                                                                                                                                                                                                                       103 ATKROKKOLEEDGIPVPSVVSRFPTAPYHHKSTNAEEREAKEPV--
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                2893 VDAEDVIGSRROPRAPKEKAOPLEDL-ASFOELSOT----
                                                                                                                                                                                               (POTENTIAL)
                                                                                                                                                                                                                                                                     3.9%; Score 137; DB 19.9%; Pred. No. 2.4;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         347 QLGSPRKEVVESAIRDLEYHDELLSLLKL----
                                                                                                                                                                                                              MISSING
 358741
                                                                                                                                                                                                                                                                                                     Conservative
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   3074 KEEHKLODSVPENK 3087
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              602 PEGPK---EVPEEE 612
                                                                                                                     2447
2569
2689
1477
1598
1720
1842
1964
2086
2204
2326
                                                                                                                                                                                                               495
                                                                                                                                                                                                                               A)
                                                                                                                                                                                                                                                                                     Best Local Similarity
Matches 146; Conserv
 1366
1487
1608
1730
1851
1974
2096
                                                                                                                     2335
2457
2579
                                                                                                                                                                 2699
                                                                                                                                                                                              NP_BIND
VARSPLIC
                                                                                                                                                                                                                             SECUENCE
                                                                                                                                                                                                                                                                        Query Match
                                                                                          REPEAT
                                                             REPEAT
                                                                                                          REPEAT
                                                                                                                      REPEAT
                                                                                                                                        REPEAT
                                                                                                                                                                   REPEAT
                                                                                                                                                                                   REPEAT
                               REPEAT
                                             REPEAT
                                                                            REPEAT
                                                                                                                                                     REPEAT
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              438
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       2968
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   147
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 qq
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           Dp
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  qq
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          οy
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       qq
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       δλ
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              qq
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                οy
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            q
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          δ
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     qq
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     δλ
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          δy
                                                                                                                                                                                                                                                                                                                                                                                              δλ
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 δλ
```

```
This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See http://www.isb-sib.ch/announce/
                                                                                                                                                                                                                                                                                                                                                                                                                                                         PTM: THERE ARE A NUMBER OF REPEATS OF THE TRIPEPTIDE K-S-P, NFM IS PHOSPHORYLATED ON A NUMBER OF THE SERINES IN THIS MOTIF. IT IS
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              OF AXONAL CALIBER.
PTM: PHOSPHORYLATION SEEMS TO PLAY A MAJOR ROLE IN THE FUNCTIONING
                                                                                                                                                                                                                                                                                                                                                                                                                                                                     PHOSPHORYLATED ON A NUMBER OF THE SERINES IN THIS MOTIF. IT IS THOUGHT THAT PHOSPHORYLATION OF NEH RESULTS IN THE FORMATION OF INTERFILAMENT CROSS BRIDGES THAT ARE IMPORTANT IN THE MAINTENANCE
                                                                                                                                                                                                                                                                                                                                                                                                                           -1- FUNCTION: NEUROFILAMENTS USUALLY CONTAIN THREE IF PROTEINS: L, M, AND H WHICH ARE INVOLVED IN THE MAINTENANCE OF NEURONAL CALIBER.
                                                                                                                                                                                                                                              Levy E., Liem R.K.H., D'Eustachió P., Cowan N.J.; "Structure and evolutionary origin of the gene encoding mouse NF-M, the middle-molecular-mass neurofilament protein."; Eur. J. Biochem. 166:71-77(1987).
                                                                                                                                           Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus. NCBI_TaxID=10090;
                                                                                                                                                                                                                                                                                                                                                             Julien J.-P., Meyer D., Flavell D., Hurst J., Grosveld F.; "Cloning and developmental expression of the murine neurofilament
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    LEVELS OF PHOSPHORYLATION BEING ALTERED DEVELOPMENTALLY AND COINCIDENT WITH A CHANGE IN THE NEUROFILAMENT FUNCTION.

-!- SIMILARITY: BELONGS TO THE INTERMEDIATE FILAMENT FAMILY.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    (BY SIMILARITY). (BY SIMILARITY).
          NPM_MOUSE STANDARD; PRT; 848 AA.
P08553; Q61961;
01-AUG-1988 (Rel. 08, Created)
01-NOV-1990 (Rel. 16, Last sequence update)
20-AUG-2001 (Rel. 40, Last annotation update)
NEUROPELIAMENT TRIPLET M PROPERIN (160 RDA NEUROFILAMENT PROTEIN)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          OF THE LARGER NEUROFILAMENT POLYPEPTIDES (NF-M AND NF-H),
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                PROSITE; PS00226; IF; 1.
Intermediate filament; Coiled coil; Neurone; Phosphorylation;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     (GLCNAC)
                                                                                                        (NEUROFILAMENT MEDIUM POLYPEPTIDE) (NF-M).
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  or send an email to license@isb-sib.ch).
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          COIL 1A.
LINKER 1.
COIL 1B.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         LINKER 12.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     O-LINKED
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          COIL 2A.
LINKER 2.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        COIL 2B.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         LINKER
                                                                                                                                                                                                                                                                                                                                  SEQUENCE OF 322-540 FROM N.A.
MEDLINE=87158637; PubMed=3103856;
                                                                                                                                                                                                                      SEQUENCE FROM N.A.
MEDLINE=87246694; Pubmed=3036526;
                                                                                                                                                                                                                                                                                                                                                                                                            Brain Res. 387:243-250(1985).
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    EMBL; X05640; CAA29127.1; -.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 M20481; AAA39815.1; -.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            filament; 1.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  IPR001664; IF.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              0
4102
8410
134
1147
1246
2263
2285
410
470
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               PIR; S00030; S00030.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 MGD; MGI:97314; Nfm.
                                                                                                                          NEF3 OR NEFM OR NFM.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          103
135
148
247
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               Pfam; PF00038;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                Glycoprotein.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    InterPro;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         CARBOHYD
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 NIT MET
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             DOMAIN
DOMAIN
DOMAIN
DOMAIN
DOMAIN
DOMAIN
DOMAIN
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         DOMAIN
DOMAIN
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     EMBL;
NFM_MOUSE
```

CARBOHYD

```
34;
                                                                                                                                                                                                                      53 TRSAVAPRLAYSSAMLSSAESSLDFSQSSSLLNGGSGGDYKLSRSNEK------EQL 103
                                                                                                                                                                                                                                                                                                    104 OGLNDRFAGYIEKVHYL---EQONKEIEAEIQALROKQASHAQLGDAYDQEIRELRATLE 160
                                                                                                                                                                                                                                                                                                                                              105 --KRQKKQ-----LEEDGIPVPSVVSRFPTAPYHHKSTNAEEREAKEPVLKTHSKDVE 155
                                                                                                                                                                                                                                                                                                                                                                                   161 MVNHEKAQVQLDSDHLEED---IHRLKERF------EEEARLRDDTEAAIRALRKDIE 209
                                                                                                                                                                                                                                                               63 AG---RMMGKDANGHCLREGKEQEEGVKMAIEGLAR-------MERRLQRAT-- 104
                                                                                                                                                                                                                                                                                                                                                                                                                              156 REAEIGVDDVVKMEPAATNII-----EPEDAQDAAERGAARPPAVNSSYLPL----- 202
                                                                                                                                                                                                                                                                                                                                                                                                                                                                  210 ESSMVKVELDKKVQSLQDEVAFLRRNHEEEVADLLAQIQASHITVERKDYLKTDISTALK 269
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  270 EIRSQLECHSDQNMHQAEEW-FKCRYAKLTEAAEQNKEAIRSAKEEIAEYRRQLQSKSIE 328
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       : |: | : | : | : | 329 LESVRGTKESLERQLSDIEERHNHDLSSYQDTIQQLENELRGTKWEMARHLREYQDLLNV 388
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            284 KMLCWNEKYGIRFLRLSSEMFPFASH-----PVHGYKL------AP---- 318
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               MASIVDHRHPLQ-----FEDEPEHHLKNKPDKSKEPQDEL-GHKFVQELGLANARDI--V 283
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     KMALDIEIAAYRKL-LEGEETRFSTFSGSITGPLYTHRQPSVTISSKIQKTKVEAPKLKV 447
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          319 ---FASEVLAEAGRVAAELGHRLTTHPGQFTQLGSPRKEVVESAIRDLEYHDELLSLLKL 375
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                448 OHKFVEEIIEET-KVEDEKSEMEETLTAIAEELAASAKEEKEEAEEKEEPEAEKSPVKS 506
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         376 PEQQNRDAVMIIHMGGQFGDKAATLERFKRNYARLSQSCKNRLVLENDDVGWTVHDLLPV 435
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               535 ......VKSDQA--EEGG-----SEKEGSSEKDEGEGEEGETEAE 568
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      Indels 249; Gaps
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                ------FSSRTCR 231
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        436 CEELNIPMVLDYHHHNICFDPAHLREGTLDISDPKLQERIANTWKRKGIKQKMHYSEPCD 495
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        496 GAVTPRDRRKHRPRVMTLPPCPPDMDLMIEAKDKEQAVFELMRTFKLPGFEKINDMVPYD 555
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       556 RDDENRPAPPVKAPK----KKKGGKRKRTTDEEAAEPEEVDTAADDVKDAPEGPK-EVP 609
                                                                                                                                                                                  3 SRKSKAAALDTPQSESSTFSSTLDSSAPSPARNLRRSGRNILQPSSEKDRDHEKRSGEEL, 62
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    ---KPKKREVKKGKVPEEVEDEG 652
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                671 VKEKAV-----KVKEKABEEG 713
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                507 PEAKEEE------EEGEKEEEEEGQEEE-------EEEDEG-----
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                611 PVEEVKPKPEAKAGKGEOKEEEKVEEEKKEVTKESPKEEKVEKKEEKPKDVADKKKAESP
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
                                                                                                Ouery Match 3.9%; Score 136.5; DB 1; Length 848; Best Local Similarity 20.1%; Pred. No. 0.47; Matches 157; Conservative 104; Mismatches 270; Indels 249,
S -> F (IN REF. 2).
QA -> RR (IN REF. 2).
; 5F251F274D0F13B6 CRC64;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        005682; Q13979; Q13978; Q14741; Q14742;
01-NOV-1997 (Rel. 35, Created)
01-NOV-1997 (Rel. 35, Last sequence update)
30-MAY-2000 (Rel. 39, Last annotation update)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               793 AA.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           610 EEERAMGGPYNRVYWPLGCEEWL-----
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             PRT:
                                          95910 MW;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         STANDARD;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           CALD1 OR CDM OR CAD.
Homo sapiens (Human).
                                      848 AA;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         CALDESMON (CDM)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         CALD_HUMAN
                                      SEQUENCE
  CONFLICT
                     CONFLICT
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                203
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               232
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       389
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                RESULT 6
CALD_HUMAN
FT
                                                                                                                                                                                                                                                                                                      qq
                                                                                                                                                                                                                                                                                                                                                                                                                                                                     q
                                                                                                                                                                                                                                                                                                                                                 ò
                                                                                                                                                                                                                                                                                                                                                                                                                            ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   q
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               qq
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            δ
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         δ
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              q
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           q
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           δλ
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         ID DT DT OC OC OC
```

```
This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See http://www.isb-sib.ch/announce/or send an email to license@isb-sib.ch).
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         IN MUSCLE
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            BINDING DOWAIN, AND THE C-TERMINAL A TROPOMYOSIN/ACTIN/CALMODULIN-BINDING DOMAIN. THESE TWO DOMAINS ARE SEPARATED BY A CENTRAL HELICAL REGION IN THE SMOOTH-MUSCLE FORM.

PTM: IN NON-MUSCLE CELLS, PHOSPHORYLATION BY CDC2 DURING MITOSIS CAUSES CALDESMON TO DISSOCIATE FROM MICROFILAMENTS.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 IS PREDOMINANTLY EXPRESSED IN SMOOTH MUSCLES, WHEREAS LOW-MOLECULAR-WEIGHT CALDESMON (L-CALDESMON) IS WIDELY DISTRIBUTED IN NON-WUSCLE TISSUES AND CELLS. NOT EXPRESSED IN SKELETAL MUSCLE OR
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     STRESS FIBERS IN FIBROBLASTS (NONMUSCLE) (BY SIMILARITY).
ALTERNATIVE PRODUCTS: THREE NON-MUSCLE ISOFORMS CALLED HELA L-CAD
I, HELA L-CAD II, WI-38 L-CAD II, WI-38 L-CAD II/1-CAD AND ONE
SMOOTH-MUSCLE ISOFORM H-CAD (SHOWN HERE) ARE PRODUCED BY
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                TISSUE SPECIFICITY: HIGH-MOLECULAR-WEIGHT CALDESMON (H-CALDESMON)
                                                                                                                                                                                   expression in normal and
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              NOMBUSCUE CELLS (COULD ACT AS A BRIDGE BETWEEN MYGSIN AND ACTIN FILAMENTS). STIMULATES ACTIN BINDING OF TROPOMYOSIN WHICH FILAMENTS STABILIZATION OF ACTIN FILAMENT STRUCTURE. IN MUSCIN SALESUES, INHIBITS THE ACTOMYOSIN ATPASE BY BINDING TO F-ACTIN. THIS INHIBITION IS ATTENUATED BY CALCIUM-CALMODULIN AND IS POTENTIATED BY TROPOMYOSIN. INTERACTS WITH ACTIN, MYOSIN, TWO MOLECULES OF TROPOMYOSIN AND MITH CALMODULIN. ALSO PLAY AN ESSENTIAL ROLE DURING CELLULAR MITHOSIS AND RECEPPOR CAPPING. SUBCELLULAR LOLAR DURING CHILAMENTS IN SMOOTH MISCLE AND ON
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 PHOSPHORYLATION REDUCES CALDESMON BINDING TO ACTIN, MYOSIN, AND
                                                                                                                                                                                                                                                                                                                                                                                                                                         MEDLINE=93101679; Pubmed=1465449;
Hayashi K., Yano H., Hashida T., Takeuchi R., Takeda O., Asada K., Takanashi E.-I., Kato I., Sobue K.;
"Genomic structure of the human caldesmon gene.";
Proc. Natl. Acad. Sci. U.S.A. 89:12122-12126(1992)
-!- FUNCTION: ACTIN- AND MYOSIN-BINDING PROTEIN IMPLICATED IN THE REGULATION OF ACTOMYOSIN INTERACTIONS IN SMOOTH MUSCLE AND
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             REORGANIZATION (BY SIMILARITY).
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            DOMAIN: THE N-TERMINAL PART SEEMS TO BE A MYOSIN/CALMODULIN-
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 CALMODULIN AS WELL AS ITS INHIBITION OF ACTOMYOSIN ATPASE ACTIVITY. PHOSPHORYLATION ALSO OCCURS IN BOTH QUIESCENT AND DIVIDING SMOOTH MUSCLE CELLS WITH SIMILAR EFFECTS ON THE
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     INTERACTION WITH ACTIN AND CALMODULIN AND ON MICROFILLAMENTS
                                                                                                                                                                                                                                                                                                                                Humphrey M.B., Herrera-Sosa H., Gonzalez G., Lee R., Bryan J.,
"Cloning of cDNAs encoding human caldesmons.",
Gene 112:197-204(1992).
                                                          SEQUENCE FROM N.A. (ISOFORM WI-38 L-CAD II).
TISSUE-Lung fibroblast:
MEDLINE-9138497; PubMed-1885618;
MEDLINE-9138497; Juh J.L.-C., Lin J.J.-C.;
"Characterization of cDNA clones encoding a human fibroblast caldesmon isoform and analysis of caldesmon expression in nor
  Eutheria; Primates; Catarrhini; Hominidae; Homo.
                                                                                                                                                                                                                                                                   SEQUENCE FROM N.A. (ISOFORMS WI-38 L-CAD II AND H-CAD)
                                                                                                                                                                                                                                                                                                                                                                                                                    SEQUENCE FROM N.A. (ISOFORMS HELA L-CAD I AND II).
                                                                                                                                                                                                                    Biol. Chem. 266:16917-16924(1991).
                                                                                                                                                                                                                                                                                      TISSUE=Aorta;
MEDLINE=9220999; PubMed=1555769;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           EMBL; M64110; AAA35636.1; -.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  . M83216; AA58420.1;
. M83216; AA58419.1;
. D90452; BAA14418.1;
. D90453; BAA14419.1;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              ALTERNATIVE SPLICING.
                                                                                                                                                                                                    transformed cells.'
                  NCBI_TaxID=9606;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           HEART
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            EMBL;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          EMBL;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      EMBL:
```

```
28;
                                                                                                                                                                                                                                                                                                                                                                                                                                                             MDDFERRRELRROKREEMRLEAER -> MLGGSGSHGRRSL
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     -----EREAE-----EREAE-320
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 166 VKMEPAATNIIEPEDAQDAAERGAARPPAVNSSYLPLPWKGRLGYACLNTYLRNAKPPIF 225
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        226 SSRTCRM-----ASIVDHRHPLQFED------EPEHHLKNKPDKSKEPQDELG 267
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 SSEEPKQEEEREQGSDEISHHEKMEEEDKERAEAERARLEAEERERIKAEQDKKIADERA 293
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             268 HKFVQELGLANARDIVKMLCWNEKYGIRFLRLSSEMFPFASHPVHGYKLAPFASEVLAEA 327
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                -----EGVKMA--IEGLARME----RRLQRATKRQKK---QLEE 113
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 114 DGIPVPS-----VVSRFPTAPYHHKSTNAEER---EAKEPVLKTHSKDVEREAEIGVDDV 165
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              188 KEKE------EEEEEKPKRG----SIGENQVEVMVEEKTTESQEETVVMSLKNGQI 233
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       68 VNAQNSVPDEEAKTTTTNTQVEGDDEAAFLERLARREERROKRLQEALERQKEFDPTITD 127
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           34 RNLRRSGRNILQPSSEKDRDHEKRSGEELAGRMMGKDANGHCLREGKEQE------ 83
                                                                        3 X 14 AA TANDEM REPEATS OF E-E-K-R-A-
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    8 RELRROKREEMRLEAERIAYORNDDDEEEAARERRRARQERLROKQEEESLGQVTDQVE 67
                                               Muscle protein; Actin-Dinding; Calmodulin-binding; Phosphorylation; Repeat; Alternative splicing.

3 x 14 AA TANDEM REPEATS OF E-E-K-R-POMAIN
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               ISOFORM WI-38 L-CAD I).
MISSING (IN ISOFORM HELA L-CAD II AND
                                                                                                                                                                                TROPOMYOSIN-BINDING (POTENTIAL).
STRONG ACTIN-BINDING (BY SIMILARITY).
WEAMODULIN-BINDING (BY SIMILARITY).
WEAK ACTIN-BINDING (BY SIMILARITY).
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  MISSING (IN ISOFORM HELA L-CAD I AND
                                                                                                                                                                                                                                                                                                                                                                                                                                                                       AALSO (IN ISOFORM HELA L-CAD I AND ISOFORM HELA L-CAD II).
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                Matches 143; Conservative 105; Mismatches 271; Indels 180;
                                                                                                                                         MYOSIN AND CALMODULIN-BINDING (BY
                                                                                                                                                                     TROPOMYOSIN-BINDING (POTENTIAL).
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        3.9%; Score 135; DB 1; Length 793;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           ISOFORM WI-38 L-CAD II/1-CAD).
                                                                                                                                                                                                                                                                                                                                                      PHOSPHORYLATION (BY CDC2) (BY
                                                                                                                                                                                                                                                                                                                                                                                  PHOSPHORYLATION (BY CDC2) (BY
                                                                                                                                                                                                                                                                                                                                                                                                           PHOSPHORYLATION (BY CDC2) (BY
                                                                                                                                                                                                                                                                                                                                                                                                                                    PHOSPHORYLATION (BY CDC2) (BY
                                                                                                                                                                                                                                                                                                                              PHOSPHORYLATION (BY CDC2) (BY
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    V -> M (IN REF. 1).
2A0DC63D16DD6B5F CRC64;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      Pred. No. 0.53;
                                                                                        A-E-E-R-Q-R-I-K
                                                                                                                                                                                                                                                                                                                                                                                                                                                  SIMILARITY)
                                                                                                                                                                                                                                                                                                                                             SIMILARITY)
                                                                                                                                                                                                                                                                                                                                                                                                SIMILARITY)
                                                                                                                                                                                                                                                                                                                                                                                                                         SIMILARITY)
                                                                                                                                                                                                                                                                                                                                                                       SIMILARITY
                                                                                                                                                                                                                                                                POLY-GLU.
                                                                                                                                                                                                                                                    POLY-THR
                                                                                                                                                                                                                                                                           -GTO
                                                                                                                                                                                                                                                                                        -ARG
                                                                                                                                                                                                                                                                                                      -GI'D
                                                                                                                                                                                                                                                                                                                   POLY-GLU
                                                                                                                                                                                                                                                                                                                                                                                                                                                                             AALSO
         InterPro; IPR000075; Caldesmon.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    93250 MW;
                      Pfam; PF02029; Caldesmon; 1.
PRINTS; PR01076; CALDESMON.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      20.5%;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      294 RIEAEEKAAAOERE----
                                                                                                                                                                                                                                                                                                                                                                                                                                      789
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     436
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              462
                                                                                                                 346
360
207
                                                                                                                                                                                  674
686
722
793
                                                                                                                                                                                                                                                                                        543
583
600
724
                                                                                                                                                                                                                                                                                                                                                        730
                                                                                                                                                                                                                                                                                                                                                                                  753
                                                                                                                                                                                                                                                                                                                                                                                                            759
                                                                                                                                                                                                                                                                                                                                                                                                                                                               24
                                                                                                                                                                     621
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     793 AA;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        Best Local Similarity
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              208
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     208
                                                                                                                                                                                                                                                                                                                                                        730
                                                                                                                                                                                                                                                                                                                                                                                                            759
                                                                                                                                                                                                                                                                                                                                                                                                                                      789
                                                                             319
                                                                                                       319
                                                                                                                                                                                                                                                                                                                                                                                   753
MIM; 114213;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           Query Match
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     VARSPLIC
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               VARSPLIC
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       SEQUENCE
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         CONFLICT
                                                                                                                                                                                                                                                                                                                                                                                                                                                                VARSPLIC
                                                                                                                                                                                                                                                                                                                                                                                                              MOD_RES
                                                                                                                                                                                                                                                                                                                                                                                                                                       MOD_RES
                                                                                                                                                                                                                                                                                                                               MOD_RES
                                                                                                                                                                                                                                                                                                                                                           MOD_RES
                                                                                                                                                                                                                                                                                                                                                                                     MOD RES
                                                                                                                                                                                                                                                                             DOMAIN
                                                                             DOMAIN
                                                                                                                                             DOMAIN
                                                                                                                                                                      DOMAIN
                                                                                                                                                                                                  DOMAIN
                                                                                                                                                                                                                          DOMAIN
                                                                                                                                                                                                                                        DOMAIN
                                                                                                                                                                                                                                                                DOMAIN
                                                                                                                                                                                                                                                                                                      DOMAIN
                                                                                                                                                                                    DOMAIN
                                                                                                       REPEAT
                                                                                                                                  REPEAT
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             pp
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 q
 qq
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     óγ
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                δλ
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        a
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   ò
```

```
This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation the European Bioinformatics Institute. There are no restrictions on its
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    chan G.K.T., Schaar B.T., Yen T.J.;
Chan G.K.T., Schaar B.T., Yen T.J.;
"Characterization of the kinetochore binding domain of CENP-E reveals
interactions with the kinetochore proteins CENP-F and hBUBRI.";
Gell Biol. 143:49-63(1998).
-!- FUNCTION: MINUS-END DIRECTED MICROTUBULE MOTOR. PROBABLE
KINETOCHORE MOTOR. ACCUMULATES JUST BEFORE MITOSIS AT THE G2 PHASE
OF THE CELL CYCLE. PROBABLY IMPORTANT FOR CHROMOSOME MOVEMENT
AND/OR SPINDLE ELONGATION.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          -!- SUBCELLULAR LOCATION: ASSOCIATES WITH KINETOCHORES DURING CONGRESSION, RELOCATES TO THE SPINDLE MIDZONE AT ANAPHASE, AND IS QUANTITATIVELY DISCARDED AT THE END OF THE CELL DIVISION.
                                                                                                                                                                                                                                                                                                                                              557 DDENRPAPPVKAPKKKKGGKRKR--TTDEE----AAEPEEVDTAADDVKDAPEGPKEVPE 610
                                                                                                                                                                                                                                                                                                                                                                         372 RARAEBEEK-AKVEEGKRN-KQLEE--KKRAMQETKIKGEKVE-----QKIEGKWVNEK 421
                                                                                                                                                                        448 HHHNICFDPAHLR-----EGT-LDISDPKLQERIANTWKRKGIKQK--MHYSEPCDGAVT 499
                                                                                                                                                                                                                 422 KAQEDKLQTAVLKKQGEEKGTKVQAKREKLQED-KPTFKKEEIKDEKIKKDKEPKEEVKS 480
                                                                                                                                                                                                                                                             500 PRDRRKHRPRVMTLPPCPPDMDLMIEAKDKEQAVFELM----RTFKLPGFEKINDMVPYDR 556
                                         321 EKRAABERORIKE----EEKRAAEERORIKEEEKRAAEERO-----RIKEEEKRAAEERO 371
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               SECUENCE FROM N.A. MEDINE-93024922; PubMed-1406971; SECUENCE-93024922; PubMed-1406971; Yen T.J., Li G., Schaar B.T., Szilak I., Cleveland D.W.; "CENP-E is a putative kinetochore motor that accumulates just before
328 GRVAABELGHRLTTHPGQFTQLGSPRKEVVESAIRDLEYHDELLSLLKLPEQQNRDAVMII 387
                                                                                  388 HMGGQFGDKAATLERFKRNYARLSQSCKNRLVLENDDVGWTVHDLLPVCEELNIPMVLDY 447
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            MEDLINE=95196755; PubMed=7889940; Thrower D.A., Jordan M.A., Schaar B.T., Yen T.J., Wilson L.; Thrower D.A., HeLa cells contain a CENP-E-associated minus end-directed
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
                                                                                                                                                                                                                                                                                                    481 FMDRKKGFTEV------KSQNGEFWTHKLKHTENTFSRPGGRASVD-----
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             -!- SUBUNIT: INTERACTS WITH CENP-F AND BUBRI KINASE.
                                                                                                                                                                                                                                                                                                                                                                                                                                        611 EERAMGGPYNRVYWPLGCEEWLKPKKREVKKGKVPEEVE 649
                                                                                                                                                                                                                                                                                                                                                                                                                                                                               581 EEEQRRKQ-----EEADRKLREEEEKRRLKEEIE 609
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    01-JUL-1993 (Rel. 26, Created)
01-JUL-1993 (Rel. 26, Last sequence update)
30-MAY-2000 (Rel. 39, Last annotation update)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             2663 AA
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                30-MAY-2000 (Rel. 39, Last annotation of CENTROMERIC PROTEIN E (CENP-E PROTEIN).
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             PRT;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             CHARACTERIZATION.
MEDLINE=98437347; PubMed=9763420;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    EMBO J. 14:918-926(1995).
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              Nature 359:536-539(1992).
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             STANDARD;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 Homo sapiens (Human)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    microtubule motor
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              CHARACTERIZATION.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               NCBI_TaxID=9606;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  CENE_HUMAN
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                mitosis
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       002224;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 CENPE.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           CENE_HUMAN
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    a
                                                                                                                                     q
                                                                                                                                                                            Óγ
                                                                                                                                                                                                                        q
                                                                                                                                                                                                                                                                      ó
                                                                                                                                                                                                                                                                                                           q
                                                                                                                                                                                                                                                                                                                                                       \overset{\circ}{\circ}
                                                                                                                                                                                                                                                                                                                                                                                                   qq
                                                                                                                                                                                                                                                                                                                                                                                                                                            Öλ
                                                                                            ζ
```

```
modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See http://www.isb-sib.ch/announce/or send an email to license@lsb-sib.ch).
                                                                                                                                                                                                                                                                                                                                                                                                                                                      32;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      --EEGVKM-----AIEGLARMERRLQRA-----TKRQKKQ 110
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            111 LEEDGIPVPSVVSRFPTAPYHHKSTNAEEREAKEPVLKTHSKDVEREAEIGVDDVVKMEP 170
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        644 LRSENLELKEKMKELATT-YKQMENDIQLYQSQLEAKKKMQVDLEKELQSAFNEITKL-- 700
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           171 AATNIIE---PEDA-------QDAAERGAARPPAVNSSYL-PLPWKGRL 208
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        460 REGTLDISDPKLQERIANTWKRKGIKQKMHYSEPCDGAVTPRDRRKHRPRVMTLPPCPPD 519
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            EENTGETKD-EFQQKMVGIDKKQDLEAKNTQTLTAD--VKDNEIIEQQRKIFSLIQEKNE 1054
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  296 FLRLSSEMFPFASHPVHGYKLAPFASEVLAEAGRVAAELGHRLTTHPGQFTQLGSPRKEV 355
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              -- CFDPAHL 459
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              847 IVNLSKEAQKFDS------SLGALKTELSYK-----TQELQEKTRE 881
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        356 VESAIRDLEYHDELLSLLKLPEQQ-NRDAVMIIHMGGQFGDKAATLERFKRNYARLSQSC 414
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                520 MDLMIEA--KDKEQAVFELMRTFK--LPGFEKINDMVPYDRDDENRPAPPVKAPKK---- 571
                                                                                                                                                                                                                                                                                                                                                                                                                                                        Gaps
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      ------SAPSPARNLRRSGRNI------- 43
 non-profit institutions as long as its content is in
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     44 -----LQPSSEKDRD-----HEKRSGEELAGRAMGKDANGHC
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     527 KLKEKNDLDEFEALERKTKKDQEMQLIHEISNLKNLVKHREVYNQDLENELSSK---VEL
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             701 --TSLIDGKVPKDLLCNLELEGKITDLQKELNKEVEENEALREEVILLSELKSLP----
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               209 GYACLNTYLRNAKPPIFSSRTCRMASIVDHRHPLQFEDEPEHHLKNKPDKSKEPQDELGH
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    882 VOERLNEMEQLKEQLENRDSPLQTVEREKTLITEKLQQTLEEVKTLTQEKDDLKQLQES-
                                                                                                                                                                                                                    PROSITE; PSO0411; KINESIN_MOTOR_DOMAIN1; 1.
PROSITE; PS50067; KINESIN_MOTOR_DOMAIN2; 1.
MOTOR protein; Cell division; ATP-binding; Colled coil; Mitosis;
                                                                                                                                                                                                                                                                                                                                                                                                                             Best Local Similarity 17.9%; Pred. No. 2.8;
Matches 143; Conservative 126; Mismatches 280; Indels 252;
                                                                                                                                                                                                                                                                                                                                                                                                                3.8%; Score 134; DB 1; Length 2663; 17.9%; Pred. No. 2.8;
                                                                                                                                                                                                                                                                                          MECHANOCHEMICAL (MOTOR).
COLLED COLL (POTENTIAL).
GLOBULAR (POTENTIAL).
ATP (BY SIMILARITY).
WW: CEFC13880C8C8C8 CRC64;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          KNRLVLENDDVGWTVHDL----LPVCEEL-NIPMVLDYHHHNI--
                                                                                                                                                                 Pfam; PF00225; kinesin; 1.
PRINTS; PR00380; KINESINHEAVY.
SMART; SM00129; KISC; 1.
                                                                                                                                                    InterPro; IPR001752; kinesin.
                                                                                EMBL; 215005; CAA78727.1; -.
                                                                                                                                                                                                                                                                                                                                                           2663 AA; 312087
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   16 SESSTFSSTLDS-----
                                                                                                                                                                                                                                                                            Cell cycle; Centromere.
                                                                                                                                                                                                                                                                                                            2471
                                                                                                                                                                                                                                                                                                                            2663
                                                                                                $28261; $28261.
; P03069; 1ZII.
                                                                                                                                                                                                                                                                                                            336
                                                                                                                                                                                                                                                                                                                                            86
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       76 LREGKEQ--
                                                                                                                                  MIM; 117143;
                                                                                                                                                                                                                                                                                                                                           NP_BIND
SEQUENCE
                                                                                                                                                                                                                                                                                                                                                                                                                Query Match
                                                                                                                                                                                                                                                                                                                              DOMAIN
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 754
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              866
ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   QQ
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       8
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         a
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             ð
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              a
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                Q
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        q
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            q
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          à
```

```
1110 KKEGELSRTCDRLAEVEEKLKEKSQQLQEKQQQLLNVQEEMSEMQKKINEI------ 1160
                                             572 KKGGKRKRTTDEEAAEPEEVDTAADDVKDAPEGPKEVPEEERAMGGPYNRVYWPLGCEEW 631
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  Submitted (MAY-1995) to the EMBL/GenBank/DDBJ databases.
-!- FUNCTION: NOT KNOWN; MULTICOPY SUPPRESSOR OF A VARIETY OF DEFECTS.
SEEMS TO INTERACTS WITH CDC42. MAY PLAY AN IMPORTANT ROLE IN CELL
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   CYCLE PROGRESSION.
-!- MISCELLANEOUS: ZDS MEANS 'ZILLION DIFFERENT SCREENS' AS BOTH ZDS1
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    Ma X.J., Lu Q., Grunstein M.; "A search for proteins that interact genetically with histone H3 and H4 amino termini uncovers novel regulators of the Swel kinase in
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       "The Saccharomyces cerevisiae SSD1 gene is involved in the tolerance to high concentration of Ca2+ with the participation of
                                                                                                                                                                                                                                                                                                         ZDS1 PROTEIN (NRC1 PROTEIN) (RT2GS1).
ZDS1 OR NRC1 OR CES1 OR CKM1 OR OSS1 OR STM2 OR HST1 OR YMR273C OR
                                                                                                                                                                                                                                                                                                                                                                          Eukaryota; Fungi; Ascomycota; Saccharomycotina; Saccharomycetes;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         "ZDS1 and ZDS2, genes whose products may regulate Cdc42p in
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       Tsuchiya E., Matsuzaki G., Kurano K., Fukuchi T., Tsukao A.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       Jia Y., Li J.-P., Butow R.A.;
Submitted (AUG-1996) to the EMBL/GenBank/DDBJ databases.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           Schwer B., Liu Y., Shuman S.;
Submitted (JUL-1995) to the EMBL/GenBank/DDBJ databases
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     STRAIN-S288C / AB972;
Lye G., Churcher C.M., Barrell B.G., Rajandream M.A.;
                                                                                                                                                                                                                                                                                                                                                                                           Saccharomycetales; Saccharomycetaceae; Saccharomyces
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       Submitted (AUG-1994) to the SWISS-PROT data bank
                                                                                                                                                                                                                                                                         (Rel. 34, Last sequence update)
(Rel. 35, Last annotation update)
                                                                                                                                                                                                                               915 AA
                                                                                                                                                                                                                                                                                                                                                             Saccharomyces cerevisiae (Baker's yeast).
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           Mol. Cell. Biol. 16:5264-5275(1996)
                                                                                                                                                                                                                               PRT;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       SEQUENCE FROM N.A. MEDLINE=96249381; PubMed=8647431;
                                                                                                                                                                                                                                                                                                                                                                                                                                                            MEDLINE=96413277; PubMed=8816439;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    MEDLINE=97075905; PubMed=8918228;
                                                                                                                                   1161 -ENLKNELTLEHMETE 1180
                                                                                                                                                                                                                         _ZDS1_YEAST STANDARD; PP50111; P41373; 01-0CT-1996 (Rel. 34, Created)
                                                                                                            632 LKPKKREVKKGKVP-EEVEDE 651
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         Saccharomyces cerevisiae.";
Genes Dev. 10:1327-1340(1996).
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             Saccharomyces cerevisiae.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       3ene 176:35-38(1996).
                                                                                                                                                                                                                                                                                                                                                                                                                                                                         , Pringle J.R.;
                                                                                                                                                                                                                                                                                                                                                                                                                                         SEQUENCE FROM N.A.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          SEQUENCE FROM N.A.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      SEQUENCE FROM N.A.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       SEQUENCE FROM N.A.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        SEQUENCE FROM N.A.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      SEQUENCE FROM N.A
                                                                                                                                                                                                                                                                                                                                                                                                            NCBI_TaxID=4932;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         STRAIN-COP161U7
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      HST1/NRC1/BFR1
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        STRAIN=RAY-3A;
                                                                                                                                                                                                                                                                           01-0CT-1996
01-NOV-1997
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        Miyakawa T.;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         Linder
                                                                                                                                                                                                               ZDS1_YEAST
                                             à
                                                                              qq
                                                                                                              ð
                                                                                                                                           Dp
```

œ

```
This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation the European Bioinformatics Institute. There are no restrictions on its way non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See http://www.isb-sib.ch/announce/
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                30:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                ----PLQPQLEGAIEIEKKEESDSE-----SLPQLQPAVSVSSTKSNSRDREEEEA--- 720
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               721 ----KKKNKKRSNTTEISNOOHSKHVQKENTDEGKAQLQAPAQEGVQTSVPVQASAPVQN 776
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     110 QLEEDGIPVPSVVSRFPTAPYHHKSTNAE-----EREAKEPVLKTHSKDVEREAEI-- 160
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           161 GVDDVVKMEPAATNIIEPEDAQDAAERGAARPPAVNSSYLPLPWKGRLGYACLNTYLRNA 220
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   479 GTTDMVK--PSAT----DDNKETKRH-----511
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                221 KPPIFSSRTCRMASIVDHRHPLQFEDEPEHHLKNKPDKSKEPQDELGHKFVQELGLANAR 280
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             512 ------QREDDNEENQGDDENEENVDS------QRELDNSK 541
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        281 DIVKMLCWNEKYGIRFLRLSSEMFPFASHPVHGYKLAPFASEVLAEAGRVAAELGHRLTT 340
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  341 HPGQFTQLGSPRKEVVESAIRDLEYHDELLSLLKLPEQQNRDAVMI-IHMGGQFGDKAAT 399
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    -----REPVEPIVLRNRPRPHRHH---- 650
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              455 DPAHLREGTLDIS-----DPKLQERIANTWKRKGIKQKMHYSEPCDGAVTPRDRRKHRPR 509
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      510 VMTLPPCPPDMDLMIEAKDKEQAVFELMRTFKLPGFE---KINDMVPYDRDDENRPAPPV 566
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              567 KAPKKKKGKR-------KRTTDEEAAE----PEEVDTAADDVKDAP--- 602
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           377 EQEKEVEQSVDEMK-------NDDEERLKLT-KNTIKVEIDPHKSPFRQQD 419
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              420 E-DSENMSSPGSIGDFQDIYNHYRQSSGEWEQEMGIEKEAEEVPVKVRNDTVEQDLELRE 478
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  542 ------KHYISLFNGGEKTEVSNKEEMNNSSTSTATSOTROKIEKTFANLFRRKPH 591
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            592 H--KHDASSSPSSSPSSS------PSIPNNDAVHVRVRKSKKLGNKSG- 631
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       400 LERFKRNYARLSQSCKNRLVLENDDVGWTVHDLLPVCEELNIPMVL-----DYHHHNICF 454
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        317 DQPQPEGHYDEGDIGFSTSQANTLDDGEFASNMPINNTMTWPERSSLRRSRFNTYRIRSQ 376
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                50 KDRDHEKRSGEELAGRMMGKDANGHCLREGKEQEEGVKMAIEGLARMERRLQRATKRQKK 109
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         Best Local Similarity 18.1%; Pred. No. 0.9;
Matches 130; Conservative 91; Mismatches 216; Indels 281; Gaps
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           12 DTPQSE-----SSTFSSTLD----SSAPS-----PAR-NLRRSGRNILQPSSE 49
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         651 ---HSRHGSQKISVKTLKDSQPQQQI------
AND ZDS2 HAVE BEEN FOUND BY A WIDE VARIETY OF GENETIC SCREENS.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      3.8%; Score 132.5; DB 1; Length 915;
                                                                                                                                                                                                                                                                                                                                                                                                                                                   94 84 S -> P (IN REF. 5).

393 393 D -> E (IN REF. 2).

413 413 S -> C (IN REF. 2).

915 AA; 103358 MW; E5F4C2117DBEIZ10 CRC64;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    .6.
                            -!- SIMILARITY: STRONG, TO YEAST ZDS2/MCS1.
                                                                                                                                                                                                         or send an email to license@isb-sib.ch).
                                                                                                                                                                                                                                                                                                                                                                                                                              POLY-GLN.
                                                                                                                                                                                                                                            EMBL; L42821; AAB38423.1; -. EMBL; U32580; AAA74626.1; -. EMBL; U63849; AAB49281.1; -.
                                                                                                                                                                                                                                                                                                                   AAB06617.1; -.
BAA08819.1; -.
CAA89256.1; -.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      18.1%;
                                                                                                                                                                                                                                                                                                                                                                                                           719
                                                                                                                                                                                                                                                                                                                                                                                      ZDS1.
                                                                                                                                                                                                                                                                                                                                                                                                         716
                                                                                                                                                                                                                                                                                                                                                                                      S0004886;
                                                                                                                                                                                                                                                                                                                      EMBL; U65682;
EMBL; D50276;
                                                                                                                                                                                                                                                                                                                                                              EMBL; Z49260;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          Query Match
                                                                                                                                                                                                                                                                                                                                                                                                                                                        CONFLICT
                                                                                                                                                                                                                                                                                                                                                                                                                                                                             CONFLICT
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   CONFLICT
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      SEQUENCE
                                                                                                                                                                                                                                                                                                                                                                                                           DOMAIN
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           Db
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     qq
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  δλ
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           Op
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          δλ
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  q
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             QΫ́
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          qq
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         δy
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               qq
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              δ
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      δy
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              Q
       δ
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           δ
```

```
the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See http://www.isb-sib.ch/announce/or send an email to license@isb-sib.ch).
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation the European Bioinformatics Institute. There are no restrictions on its
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 -!- SUBUNIT: HOMODIMER OR HOMOTETRAMER.
-!- DOMAIN: THE N-TERMINUS INTERACTS WITH ACTIN, THE C-TERMINUS WITH
VIMENTIN, DESMIN, GFAP, CYTOKERATINS, LAMIN B; WHEREAS BOTH THE N-
AND THE C-TERMINUS CAN BIND INTEGRIN BETA-4.
-!- PTM: PHOSPHORYLATED BY CDC2; REGULATES DISSOCIATION FROM
INTERMEDIATE FILLAMENTS DURING MITOSIS.
603 -----EGPKEVPEEERAMGGPYNRVYWPLGCEEWLKPKKREVKKGKVPEEVEDEGEF 654
                                    777 SAPVQTSAFVEASAQTQAPAAP-----PLKHTSILPPRKLTFADVKKPDKPNSPVQF 828
                                                                                                                                                                                                                                                                                                                                                                                                                                                          "The 300-kDa intermediate filament-associated protein (IFAP300) is namster plectin ortholog.";
                                                                                                                                                                                                                                                                                            Enkaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Cricetinae;
                                                                                                                                                                20-AUG-2001 (Rel. 40, Created)
20-AUG-2001 (Rel. 40, Last sequence update)
20-AUG-2001 (Rel. 40, Last annotation update)
PHECTIN 1 (PLTN) (PCN) (300-KDA INTERMEDIATE FILAMENT-ASSOCIATED
                                                                                                                                                                                                                                                                                                                                                                                                      MEDLINE=20334248; PubMed=10873583;
Clubb B.H., Chou Y.-H., Herrmann H., Svitkina T.M., Borisy G.G.,
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          -:- SIMILARITY: CONTAINS 1 ACTIN-BINDING DOMAIN.
-:- SIMILARITY: CONTAINS 2 CALPONIN-HOMOLOGY DOMAINS (CH).
-:- SIMILARITY: CONTAINS 3 PLECTIN REPEATS.
-:- SIMILARITY: CONTAINS 4 SPECTRIN REPEATS.
-:- SIMILARITY: BELONGS TO THE PLAKIN OR CYTOLINKER FAMILY.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             Biochem. Biophys. Res. Commun. 273:183-187(2000).
                                                                                                                               PRT; 4473 AA.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     PROSITE, PS00019; ACTININ_1; PARTIAL.
PROSITE; PS00020; ACTININ_2; FALSE_NEG.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          Interpro; IPR001589; Actinin_act_bind.
Interpro; IPR001715; Calponin_hom.
                                                                                                                                                                                                                                                                            Cricetulus griseus (Chinese hamster).
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         InterPro; IPR001101; Plectin_repeat.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            Pfam; PF00307; CH; 2.
Pfam; PF00681; Plectin_repeat; 20.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     MEDLINE=96215219; PubMed=8626512;
                                                                                                                                                                                                                                          PROTEIN) (IFAP300) (FRAGMENT).
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             InterPro; IPR002017; Spectrin.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   EMBL; AF260753; AAF70372.1; -.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               SMART; SM00033; CH; 2.
SMART; SM00250; PLEC; 32.
SMART; SM00150; SPEC; 4.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                hamster plectin ortholog.
                                                                                                                                 STANDARD;
                                                                                                                                                                                                                                                                                                                                                                                           SEQUENCE FROM N.A.
                                                                                                                                                                                                                                                                                                                                                  NCBI_TaxID=10029;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      PHOSPHORYLATION.
                                                                                                                                                                                                                                                                                                                                                                                                                                             Goldman R.D.;
                                                                                                                               PLE1_CRIGR
Q9JI55;
                                                                                                                                                                                                                                                                                                                                    Cricetulus.
                                                                                                                PLE1_CRIGR
                                                                                              RESULT
```

```
qq
                                                                               \dot{o}
                                                                                                            q
                                                                                                                                           à
                                                                                                                                                                        DP
DP
                                                                                                                                                                                                         Οy
                                                                                                                                                                                                                                      qq
                                                                                                                                                                                                                                                                   οy
                                                                                                                                                                                                                                                                                                  q
                                                                                                                                                                                                                                                                                                                                 δà
                                                                                                                                                                                                                                                                                                                                                            QQ
                                                                                                                                                                                                                                                                                                                                                                                                                          qq
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      Db
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                          ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      δ
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      2049 EQAQARAQAQAAAEKLRKEAEQEAARRAQAEQAALKQK-------QAADAEMEKH 2096
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            1997 GEOSMLERLRGEAEAARRAAEEAEEARE------QAEREAAQSRKQVEEAERLKQSAE 2048
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     1878 RSSAEDTMRSKEQAEQEAARQRQLAAEEEQRRREAEERVQKSLAAEEEAARQRKAALEEV 1937
              Coiled coil; Repeat; Structural protein; Cytoskeleton; Actin-binding;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              ------PAATNIIEPEDAQDAAERGAARPPAVNSSYLPLPWKGRLGYACLNTYLRNA 220
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    64 GRMMGKDANGHCLREGKEQEEGVKMAIEGLARMERRLQRATK-----RQKKQLEEDGI 116
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           Query Match 3.8%; Score 131.5; DB 1; Length 4473; Best Local Similarity 18.0%; Pred. No. 7.8; Matches 138; Conservative 120; Mismatches 322; Indels 187; Gaps
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        7 KAAALDTPQS-ESSTFSSTLDSSAPSPARNLRRSGRNILQPS--SEKDRDHEKRSGEELA 63
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    BINDING TO INTERMEDIATE FILAMENTS (BY
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                117 PVPSVVSRFPTAPYHHKSTNAEEREAKEPVLKTHSKDVEREAEIGVDDVVKME-----
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   SIMILARITY).
4 X 4 AA TANDEM REPEATS OF G-S-R-X.
PHOSPHORYLATION (BY CDC2).
                                                                             CENTRAL FIBROUS ROD DOMAIN
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                E144615D361E3484 CRC64;
                                                                                                                                                                                     SPECTRIN 3.
SPECTRIN 4.
COILED COIL (POTENTIAL).
                                                                                          GLOBULAR 2.
ACTIN-BINDING
                                                             GLOBULAR 1.
                                                                                                                                                        SPECTRIN 1.
SPECTRIN 2.
                                                                                                                                                                                                                                    PLECTIN 1.
                                                                                                                                                                                                                                                                                             PLECTIN 5.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       PLECTIN 32
                                                                                                                                                                                                                                                 PLECTIN 2
                                                                                                                                                                                                                                                                                                                                                                                        PLECTIN DELECTIN DELE
                                                                                                                                                                                                                                                                                                             PLECTIN PLECTIN
                                                                                                                                                                                                                                                                                                                                                         PLECTIN
PLECTIN
                                                                                                                                                                                                                                                                                                                                                                                                                                                                  PLECTIN
PLECTIN
PLECTIN
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             PLECTIN
                                                                                                                                                                                                                                                                   PLECTIN
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      PLECTIN
                                                                                                                                                                                                                                                                                PLECTIN
                                                                                                                                                                                                                                                                                                                                            PLECTIN
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               PLECTIN
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            PLECTIN
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     PLECTIN
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                PLECTIN
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              PLECTIN
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             PLECTIN
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           PLECTIN
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          PLECTIN
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         PLECTIN
                                                                                                                                                                                                                                                                                                                                                                                                                                                       PLECTIN
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  PLECTIN
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  4328
AA; 509015
PROSITE; PS50021; CH; 2
                                                                                                                                                                                                                 2548
2652
2690
2728
2766
2942
                                                                                                                                                                                                                                                                                                                                                                                                      3349
                                                                                                                                                                                                                                                                                                                                                                                                                                     3425
3463
3646
3684
3722
3760
3797
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               3834
3889
3927
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          3965
                                                                                                                                                                                                                                                                                                                          2980
3018
3056
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         4041
4094
4234
                                                                                                                                                                                                                                                                                                                                                                        311
                           Phosphorylation
NON_TER
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                MOD_RES
SEQUENCE
                                                                          DOMAIN
                                                                                          DOMAIN
                                                                                                                                                                                    REPEAT
                                                                                                                                                                                                                                                                                                                                                                                                      REPEAT
REPEAT
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                REPEAT
REPEAT
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            REPEAT
REPEAT
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          REPEAT
REPEAT
                                                                                                         DOMAIN
                                                                                                                                                                                                                                                                                                                                                         REPEAT
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  DOMAIN
                                                                                                                                        DOMAIN
                                                                                                                                                      REPEAT
                                                                                                                                                                       REPEAT
                                                                                                                                                                                                     REPEAT
                                                                                                                                                                                                                  DOMAIN
                                                                                                                                                                                                                                    REPEAT
                                                                                                                                                                                                                                                 REPEAT
                                                                                                                                                                                                                                                                REPEAT
                                                                                                                                                                                                                                                                                REPEAT
                                                                                                                                                                                                                                                                                               REPEAT
                                                                                                                                                                                                                                                                                                             REPEAT
                                                                                                                                                                                                                                                                                                                            RPEAT
                                                                                                                                                                                                                                                                                                                                          REPEAT
                                                                                                                                                                                                                                                                                                                                                                        REPEAT
                                                                                                                                                                                                                                                                                                                                                                                        REPEAT
                                                                                                                                                                                                                                                                                                                                                                                                                                     REPEAT
                                                                                                                                                                                                                                                                                                                                                                                                                                                       REPEAT
                                                                                                                                                                                                                                                                                                                                                                                                                                                                    REPEAT
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   REPEAT
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   REPEAT
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                REPEAT
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           REPEAT
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        REPEAT
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          REPEAT
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        REPEAT
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       REPEAT
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      REPEAT
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      REPEAT
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    DOMAIN
qq
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                ð
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            ò
```

```
2215 EAARLRQLAEEDLAQQRALAEKMLKEKMQAVQEATRLKAE------AELLQQQK 2262
                                                                                                                                                                                                                                                                                                                                                                                                                                                                     2097 KK--FAEQTLROKAQVEQELTTLRLQLEETDHQKSILDEELQRLKAEVTEAARQRSQVEE 2154
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               2322 QARAEEDAQRFRKQAEEIGEKLHRTELATQEKVTLVQTLEIQRQQSDHDAERLREAIAEL 2381
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          2382 EREK--EKLKQEAKLLQLKSEEMQTVQOEQILQETQALQKSFLSE------KDSLLQR 2431
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         2469 QEKQELVASMEEARRRQCEAEEAVRRKQEELQHLELQRQQQEKLLAEENQRLRERLQRLE 2528
KPPIFSSRTCRMASIVDHR-------HPLQFEDEPEHHLKNKPDKSKEPQDELGH 268
                                                                                                                                                                                                                                                                                  312 HCYKLAPFASEVLA------EAGRVAAELGHRLTTHPGQFTQLGSPRK 353
                                                                                                                                                                                                                                                                                                                                                                                                                         354 EVVESAIRDLEYHDELLSLLKLPEQQNRDAVMIIHMGGQFGDKAATLERFKRNYARLS-- 411
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 412 ---- OSCKNRLVLENDDVGWTVHDL-LPVCEELNIPMVLDYHHHNICFDPAHLREGTLDI 466
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           568 APKKK-----KGGKRKRTTDEEAA--EPEEVD-----TAADDVKDAPEGPKEVP 609
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        SDPKLQERIANTWKRKGIK-------QKMHYSEPCDGAVTPRDRRKHR 507
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   508 PRVMTLPPCPPDMDLMIEAKDKEQAVFELMRTFKLPGFEKINDMVPYDRDDENRPAPPVK 567
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           Kelly B.M.; Gillespie C.S., Sherman D.L., Brophy P.J.; "Schwann cells of the myelin-forming phenotype express neurofilament protein NF-M.";
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         Napolitano E.W., Chin S.S.M., Colman D.R., Liem R.K.H.; "Complete amino acid sequence and in vitro expression of rat NF-M, the middle molecular weight neurofilament protein."; J. Neurosci. 7:2590-259(1987).
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              MEDLINE=92165797; PubMed=1537832; Xu Z.-S., Liu W.-S., Willard M.B.; "Identification of six phosphorylation sites in the COOH-terminal tail region of the rat neurofilament protein M.";
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       01-0CT-1989 (Rel. 12, Created)
01-MAY-1992 (Rel. 22, Last sequence update)
01-MAY-1992 (Rel. 40, Last annotation update)
20-AUG-2001 (Rel. 40, Last annotation update)
NEUROFILAMENT TRIPLET M PROTEIN (160 KDA NEUROFILAMENT PROTEIN)
(NEUROFILAMENT MEDIUM POLYPEPTIDE) (NF-M).
                                                                                                                                       KF-----VQELGLANAR---DIVKMLCWNEKYGIRFLRLSSEMFPFASHPV-
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              || || || : |
2529 EEHRAALAHSEEIAATQAAAAKALPNGRDALDGP-SMEVEPEHAFEG 2574
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  610 EEERAMGGPYNRVYWPLGCEEWLKPKKREVKKGKVPEEVEDEGEFDG 656
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    845 AA
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        PHOSPHORYLATION SITES, AND REVISION TO 500.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    PRT;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        MEDLINE=87282618; PubMed=2441012;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               Cell Biol. 118:397-410(1992).
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                STANDARD;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      Rattus norvegicus (Rat).
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          NEF3 OR NEFM OR NFM.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       SEQUENCE FROM N.A.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  SEQUENCE FROM N.A.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     2432 ERFI-----
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  NCBI_TaxID=10116;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                NFM_RAT
P12839; Q63370;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          STRAIN=WISTAR;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           10
                                                                                                                                           269
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        467
```

```
This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See http://www.isb-sib.ch/announce/
                                                                                                                L and M.",

J. Biol. Chem. 268:16679-16687(1993).

J. Biol. Chem. 268:16679-16687(1993).

J. Biol. Chem. PURDEILAMENTS USUALLY CONTAIN THREE IF PROTEINS: L, M, I. FUNCTION: NEUROFILAMENTS USUALLY CONTAIN THREE IF PROTEINS: L, M, AND H WHICH ARE INVOLVED IN THE MAINTERNANCE OF NEURONAL CALLBER.

J. PTM: THERE ARE A NUMBER OF REPEATS OF THE TRIPEPTIDE K-S-P, NFM IS PHOSPHORYLATED ON A NUMBER OF THE SERIRES IN THIS MOTIF. IT IS THOUGHT THAT PHOSPHORYLATION OF NFH RESULTS IN THE FORMATION OF INTERFILAMENT CROSS BRIDGES THAT ARE IMPORTANT IN THE MAINTENANCE OF AXONAL CALLBER.
                                                                                                                                                                                                                                            PTM: PHOSPHORYLATION SEEMS TO PLAY A MAJOR ROLE IN THE FUNCTIONING THE LARGER NEUROFILAMENT POLYPEPTIDES (NF-M AND NF-H), THE LEVELS OF PHOSPHORYLATION BEING ALTERED DEVELOPMENTALLY AND COINCIDENT WITH A CHANGE IN THE NEUROFILAMENT FUNCTION, SIMILARITY: BELONGS TO THE INTERMEDIATE FILAMENT FAMILY.
                                                                                  'Glycosylation of mammalian neurofilaments. Localization of multiple
                                                                                                   O-linked N-acetylglucosamine moieties on neurofilament polypeptides
                                          MEDLINE=93346421; PubMed=8344946;
Dong D.L.-Y., Xu Z.-S., Chevrier M.R., Cotter R.J., Cleveland D.W.,
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       PROSITE; PS00226; IF; 1. Intermediate filament; Coiled coil; Neurone; Phosphorylation;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       R -> P (IN REF. 2).
V -> L (IN REF. 2).
MISSING (IN REF. 1).
316C41655B11197D CRC64;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          PHOSPHORYLATION.
MISSING (IN REF. 2).
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         O-LINKED (GLCNAC).
/FTId=CAR_000131.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             O-LINKED (GLCNAC).
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           FTId=CAR_000130.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      PHOSPHORYLATION.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     PHOSPHORY LATION.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                PHOSPHORYLATION.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              PHOSPHORYLATION.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   PHOSPHORYLATION
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  BY SIMILARITY HEAD.
                                                                                                                                                                                                                                                                                                                                                                                                                     send an email to license@isb-sib.ch).
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       COIL 1B.
LINKER 12.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 COIL 2A.
LINKER 2.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           TAIL.
COIL 1A.
LINKER 1.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               COIL 2B.
J. Biol. Chem. 267:4467-4471(1992).
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                ROD
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    95660 MW:
                                                                                                                                                                                                                                                                                                                                                                                                                                                  EMBL; M18628; AAA41696.1; -.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                            GlycoSuiteDB; P12839; -.
InterPro; IPR001664; IF.
Pfam; PF00038; filament; 1.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                  CAA78136.1; -
                                CARBOHYDRATE-LINKAGE SITES
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      147
246
263
285
289
410
47
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         430
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  506
603
608
608
666
17
21
204
500
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  845 AA;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            411
103
135
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   0
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 264
286
290
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     Glycoprotein.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                  EMBL; Z12152;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     CONFLICT
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     INIT_MET
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               CARBOHYD
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            CARBOHYD
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     MOD_RES
MOD_RES
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           CONFLICT
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         CONFLICT
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  SEQUENCE
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  MOD_RES
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              MOD_RES
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            MOD_RES
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       DOMAIN
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                DOMAIN
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              DOMAIN
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          DOMAIN
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        DOMAIN
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    DOMAIN
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  DOMAIN
```

```
Query Match
3.8%; Score 131; DB 1; Length 845;
Best Local Similarity 20.3%; Pred. No. 1;
Matches 155; Conservative 111; Mismatches 276; Indels 222; Gaps 37;
4 RKSKAAALDTPQSESSTFSSTLDSSAPSPARNLRRSGRNILQPSSEKDRDHEKRSGEELA 63
```

<u>:</u>

_ _ _ :: - :: :: ::

ŏ

-----QFTQLGSPRKEV----VESAIRDLEYHDELLSLLKLPEQQNRDAVMII--HMG 390 391 GQFGDKAATLERFKRNYARLSQSCKNRLVLENDDVGWTVHDLLPVCEELNIPMVLDYHHH 450 451 NICEDPAHLREGTLDISDPKLQERIANTWKRKGIKQKMHYSEPCDGAVTPRDRRKHRPRV 510 569 --PKKKKGGKRKRTTDEEAAEPEEVDTAADDVKDAPEGPK-EVPEEERAMGGPYNRVYWP 625 162 VNHEKAQVQIDSDHLEED---IHRLKERF-----EEEARLRDDTEAAIRAVRKDIEE 210 ------FSSRTCRM 232 271 IRSQLECHSDQNMHQAEEW-FKCRYAKLTEAAEQNKEAIRSAKEEIAEYRRQLQSKSIEL 329 233 ASIVDHRHPLQ-----FEDEPEHHLKNKPDKSKEPQDEL-GHKFVQELGLANARDI--VK 284 330 ESVRGTKESLERQLSDIEERHNHDLSSYQDTIQQLENELRGTKWEMARHLREYQDLLNVK 389 285 MICWNEKYGIRFLRISSEMFPFASHPVHGYKLAPFASEVLAEAGRVAAEL-GHRLTTHPG 343 390 MALDIEIAAYRKL-LEGEETRFSTF-------SGSITGPLYTHR---QPS 428 429 VIISSKIQKTKVEAPKLKVQHKFVEEIIEETKVEDE------KSEMEDALTVIAEELA 480 511 MTLPPCPPUMDLMIEAKDKEQAVFELMRTFKLPGFEKINDMVPYDRDDENRPAPPVKA-- 568 ---KTEGKVEEMAIKELIKVEK-P--EKAKSPVPKSPVEEVKPKPEAKAGK 623 Borrow J., Stanton V.P. Jr., Andresen J.M., Becher R., Behm F.G., Chaganti R.S.K., Civin C.I., Disteche C., Dube I., Frischauf A.M., Horsman D., Mitelman F., Volinia S., Watmore A.E., Housman D.E., "The translocation t(8:16)(pll;pl3) of acute myeloid leukaemia fuses a putative acctyltransferase to the CREB-binding protein."; 54 RSALAPRLAYSSAMLSSAESSLDFSQSSSLLNGGSGGDYKLSRSNEK------EQLQ 104 ---- MERRLQRAT --- 104 105 GLNDRFAGYIEKVHYL---EOONKEIEAEIHALROKQASHAQLGDAYDQEIRELRATLEM 161 105 -KRQKKQ-----LEEDGIPVPSVVSRFPTAPYHHKSTNAEEREAKEPVLKTHSKDVER 156 EAEIGVDDVVKMEPAATNII-----EPEDAQDAAERGAARPPAVNSSYLPL-----211 SSMVKVELDKKVQSLQDEVAFLRSNHEEEVADLLAQIQASHITVERKDYLKTDISTALKE 533 CVKSDOA--EEGG-----SEKEGSSEKDEGEGEEGETE-AEGEGEEAEAKEEK---624 DEQKEEEKVEEKKEVAKESPKEEKVEKKEEKPKDVPDKKKAESPVKEKAV-----Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo MONOCYTIC LEUKEMIA ZINC FINGER PROTEIN (ZINC FINGER PROTEIN 220). ZNF220 OR MOZ. ------KPKKREVKKGKVPEEVEDEG 652 674 ---EEMITITKSVKVSLEKDTKEEKPQQQE----KVKEKAEEEG 710 64 G---RMMGKDANGHCLREGKEQEEGVKMAIEGLAR----PRT; 2004 AA 15-JUL-1998 (Rel. 36, Created) 15-JUL-1998 (Rel. 36, Last sequence update) 20-AUG-2001 (Rel. 40, Last annotation update MEDLINE=96376968; PubMed=8782817; STANDARD; Homo sapiens (Human). SEOUENCE FROM N.A. 626 LGCEEWL----NCBI_TaxID=9606; MOZ_HUMAN 092794; RESULT 11 MOZ_HUMAN q ŏλ qq Ω ŏλ qq Öζ Op οy QQ õ q $^{\circ}$ qu δ qq ΩŸ

```
between the Swiss Institute of Bioinformatics and the EMBL outstation the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial
                                                                                                                      This SWISS-PROT entry is copyright. It is produced through a collaboration
                                                                                                                                                                                                entities requires a license agreement (See http://www.isb-sib.ch/announce/
or send an email to license@isb-sib.ch).
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           32;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                -----SSSPPILTKPTLKRK-KPF----LHRR 1020
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             787 SEEEEEEAEEGENEEPQCQERELEISVGKSVS----HENKEQD---SYSVESEKKPEVMA 839
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           101 QRATKRQKKQ-LEEDGIPVPSVVSRFPTAPYHHKSTNAEER----EAKEPVLKTHSKDVE 155
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        156 REAEIGVDDVVKMEPAATNIIEPEDAQDAAERGAARPPAVNSSYLPL------PWKGR 207
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   898 QYGECG----EKSEATQEQYTESEEQLVASEE---QPSQDGKPDLPKRRLSEGVEPWRGQ 950
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   208 LGYACLNTYLRNAKPPIFSSRTCRMASIVDHRHPLQFEDEPEHHLKNKPDKSKEPQDELG 267
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              268 HKFVQELGLANARDIVKMLCWNEKYGIRFLRLSSEMFPFASHPVHGYKLAPFASEVLAEA 327
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           328 GRVAAELGHRLTTHPGQFTQLGSPRKEVVESAIRDLEYHDELLSL----LKLPEQQNRDAV 384
-!- FUNCTION: MAY REPRESENT A CHROMATIN-ASSOCIATED ACETYLTRANSFERASE.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 SSEKDRDHEKRSGEELAGR-----MMGKDANGHCLREGKEQEEGVKMAIEGLARMERRL 100
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           Gaps
         SUBCELLUIAR LOCATION: NUCLEAR.
DISEASE: PARTICIPATES IN A T(8;16)(P11;P13) CHROMOSOMAL
TRANSLOCATION THAT PRODUCES A MOZ-CBP CHIMAERA OBSERVED IN THE
M4/M5 SUBTYPE OF ACUTE MYELOID LEGKEMIA.
SIMILARITY: CONTAINS 2 PHD-TYPE ZINC FINCERS.
SIMILARITY: BELONGS TO THE MYST (SAS/MOZ) FAMILY.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         840 PVSSTRLSKQVLPHDSLPANSQPSR--RGRWGRKNRKTQERFGDKDSKLLLEETSSAPQE
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                BREAKPOINT FOR TRANSLOCATION TO FORM
                                                                                                                                                                                                                                                                                                                                                                                                    translocation; Zinc-finger; Repeat;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           DB 1; Length 2004;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             9FFBBAC3792854BA CRC64;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      3.7%; Score 128.5; DB 1;
20.1%; Pred. No. 4.3;
Live 89; Mismatches 260;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 POLY-SER.
GLN/PRO-RICH.
MET-RICH.
                                                                                                                                                                                                                                                                                                                                                                                                                                 PHD-TYPE 1. PHD-TYPE 2.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                              -TYPE.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         POLY-GLU.
POLY-ARG.
POLY-GLU.
POLY-LYS.
GLU-RICH.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     GLU-RICH.
POLY-GLU.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              -GLU.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                POLY - SER.
                                                                                                                                                                                                                                                                                      InterPro; IPR001386; Linker_histone.
InterPro; IPR002717; MOZ_SAS.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           2004 AA; 225054 MW;
                                                                                                                                                                                                                                         EMBL; U47742; AAC50662.1; -.
                                                                                                                                                                                                                                                                                                                                                                                                 Proto-oncogene; Chromosomal
                                                                                                                                                                                                                                                                                                                      InterPro; IPR001965; PHD.—
Pfam; PF01853; MOZ_SAS; 1.
Pfam; PF00628; PHD; 2.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      Conservative
                                                                                                                                                                                                                                                                                                                                                                    SMART; SM00526; H15; 1.
SMART; SM00249; PHD; 2.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        1078
1150
1242
1302
1414
1597
1704
1977
                                                                                                                                                                                                                                                            Q60631; 4GBQ.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      Local Similarity
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                1643
1897
1546
                                                                                                                                                                                                                                                                                                                                                                                                                 Nuclear protein
                                                                                                                                                                                                                                                                           601408;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   Matches 136;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           SEQUENCE
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        Query Match
                                                                                                                                                                                                                                                                                                                                                                                                                                              ZN_FING
DOMAIN
                                                                                                                                                                                                                                                                                                                                                                                                                                                                              ZN_FING
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                PR-
                                                                                                                                                                                                                                                                                                                                                                                                                                 ZN_FING
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      DOMAIN
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 DOMAIN
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              166
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 47
 g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              QQ
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             g
```

```
This SWISS-PROT entry is copyright. It is produced through a collaboration
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          between the Swiss Institute of Bioinformatics and the EMBL outstation—
the European Bioinformatics Institute. There are no restrictions on its
use by non-profit institutions as long as its content is in no way
modified and this statement is not removed. Usage by and for commercial
entities requires a license agreement (See http://www.isb-sib.ch/announce/
or send an email to license@isb-sib.ch).
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              1122 LKPVSLLRKRDVKNSPLEPDTSTPLKKKKGW-----PKGKSRKPIHWKKRPGRKPGFKLS 1176
1021 RRVRKRKHHINSSV----VTETISETTEVLDEPFEDSDSERPMPRLEPTFEIDEBEBEE-- 1074
                                                                                                                                                     ---DEN---ELFPREYFRRLSSQDVLRCQSSSKRKSKDEEEDEESDDADDTPI 1121
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  ----S 1326
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   OF AXONAL CALIBER.
PTM: PHOSPHORYLATION SEEMS TO PLAY A MAJOR ROLE IN THE FUNCTIONING
                                                                                                                                                                                                                                                                                                                                                                                                                                                    1177 REIMPVSTQACVIEPI--VSIPKAGRK--PKIQESEETVEP-KEDMPLPEERKEEEEMQA 1231
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 1232 EAEEAEEGEEEDAASSEVPAASPADSSNSPETETKEPEVEEEEEKPRVSEEOROSEEEOO 1291
                                                                          385 MIIHMGGQFGDKAATLERFKRNY-----ARLSQSCKNRLVLENDDVGWTVHDLLPV 435
                                                                                                                                                                                                                            436 CEELNIPMVLDYHHHNICFD---PAHLREGTLDISDPKLQERIANTW-KRKGIKQKMHYS 491
                                                                                                                                                                                                                                                                                                                                                                                 ---EPCDGAVTPRDRRKHRPRVM----TLPPCPPDMDLMIEAKDKEQAVF 534
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                --EAAEPEEVDTAA-----DDVKDAPEGPKEVPEEERAMGGPYNRVYWPLGCEEWLK 633
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           Lees J.F., Shneidman P.S., Skuntz S.F., Carden M.J., Lazzarini R.A., "The structure and organization of the human heavy neurofilament
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              ELMRTFKLPGFEKINDMVP--YDRDDENRPAPPVKAPKKKKGGKRKRTTDE-----
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   Eukaryota, Metazoa, Chordata, Craniata, Vertebrata, Euteleostomi,
Mammalia, Eutheria, Primates, Catarrhini, Hominidae, Homo.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          OF THE LARGER NEUROPILAMENT POLYPEPTIBES (NF-M AND NF-H), TH
LEVELS OF PHOSPHORYLATION BEING ALTERED BEVELOPMENTALLY AND
COINCIDENT WITH A CHANGE IN THE NEUROFILAMENT FUNCTION.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   01-0CT-1989 (Rel. 12, Created)
01-0CT-1989 (Rel. 12, Last sequence update)
20-AUG-2001 (Rel. 40, Last annotation update)
NEUROFILAMENT TRIPLET H PROTEIN (200 KDA NEUROFILAMENT PROTEIN)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                SIMILARITY: BELONGS TO THE INTERMEDIATE FILAMENT FAMILY.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       1292 ELEEPEPEEEEDAAAETAQNDDHDADDEDDGHLE----------
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             (NEUROFILAMENT HEAVY POLYPEPTIDE) (NF-H).
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   MEDLINE=88328981; PubMed=3138108;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        1327 TKKKELEEQPTREDVKEE 1344
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 634 PKKREVKKGKVPEEVEDE 651
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 STANDARD;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           sapiens (Human)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        SEQUENCE FROM N.A.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     NCBI_TaxID=9606;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    NEFH OR NFH.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     NFH_HUMAN
                                                                                                                                                                                                                                                                                                                                                                             492
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            584
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 NFH_HUMAN
                                                                          öλ
                                                                                                                                                 qq
                                                                                                                                                                                                                            ó
                                                                                                                                                                                                                                                                                              Db
                                                                                                                                                                                                                                                                                                                                                                         δ
                                                                                                                                                                                                                                                                                                                                                                                                                                                qq
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              δλ
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 Db
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            Qγ
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 οy
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 11D
D111D
D121D
D131D
D1
```

EMBL; X15306; CAA33366.1; -.

```
30;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              718 SPEKAKSPVKEEAKTPEKAKS------E 748
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             PAHLRE--GILDISDPKLQERIANTWKRKGIKQKMHYSEPCDGAVTPRDRKHRPRVMTL 513
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           514 PPCPPDMDLMIEAKDKEQAVFELMRTFKLPGFEKINDMVPYDRDDENRPAPPVKAPKKKK 573
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                801 SPLKADAKAPEKEIPKKEEVKSPVKEEEKPQEVKVKE-PPKKAEEEKAPATP--KTEEKK 857
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       221 KPPIFSSRTCRMASIVDHRHPLQFEDEPEHHLKNKPDKSKEPQDELGHKFVQELGLANAR 280
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    281 DIVKMLCWNEKYGIRFLRLSSEMFPFASHPVHGYKLAP--FASEVLAEAGRVAAELGHRL 338
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  339 TTHPGOFTQLGSPRKEVVESAIRDLEYHDELLSLLKLPFQQNRDAVMIIHMGGQFGDKAA 398
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                689 ------SPVKEEAK 717
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              TLERFY --- RNYARLSQSCKNRLVLENDDVGWTVHDLLPVCEELNIPMVLDYHHHNICFD 455
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       574 GGKRKRTTDEEAAEP---EEVDTAADDVKDAP-EGPKEVPEEERAMGGPYNRVYWPLGCE 629
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    858 DSKKEEAPKKEAPKVEEKKEPAVEKPKESKVEAKKEEAEDKKKVPTPEKEAPAKVEVK 917
                                                                                                                                                                                                                                                                                                                                                                                                                                   105 KRQKKQLEEDGIPVPSVVSRFPTAPYHHKSTNAEEREAKEPV-LKTHSKDVEREAEIGVD 163
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          508 EAASPEKE----AKSPVKEEAKSPAEAKS--PEKEEAKSPAEVKSPEK----AKSPAK 555
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         164 DVVKMEPAATNIIEPEDAQDAAERGA---ARPPAVNSSYLPLPWKGRLGYACLNTYLRNA 220
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      556 EEAKSPPEAKS-PEKEEAKSPAEVKSPEKAKSPAKEEAKSPAEAKSP-----EKA 604
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    605 KSPVKE----EAKSPAEAKSPVKEEAKSPAEVKS-PEKAKSPTKEEAKS--PEKAKSPEK 657
                                                                                                                                                                                                                                                                                                                                                                                                               45 QPSSEKDRDHEKRSGEELAGRMMGKDANGHCLREGKEQEEGVKMAIEGLARMERRLQRAT 104
                                                                                                                                                                                                                                                                                                                                                                                  Gaps
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            749 KAKSPEKAKTLDVKSPEAKTPAKEEARSPADKFPEKAKSPVKEEVKSPEKAK-----
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                EEAK------SPEK---AKSPVKAEAKSPEKAKSPVKAEA------
                                                                                                                                                                                                                                                                                                                                                                                 84; Mismatches 243; Indels 164;
                                                                                                        PROSITE; PS00226; IF; 1.
Intermediate filament; Coiled coil; Neurone; Phosphorylation.
DOMAIN 1 100 HEAD.
                                                                                                                                                                                                                                                                                                                                                 3.7%; Score 128; DB 1; Length 1020;
                                                                                                                                                                                                                                                                                      13 COIL 2B.
111780 MW; 1177C9DCB3DCF1D4 CRC64;
                                                                                                                                                                                                                                                                                                                                                                  Pred. No. 1.9;
                                                                                                                                                                                                                     COIL 1B.
LINKER 12.
                                                                                                                                                                                                                                                      COIL 2A.
LINKER 2.
                                                                                                                                                                         TAIL.
COIL 1A.
LINKER 1.
             EMBL; X15308; CAA33366.1; JOINED.
EMBL; X15309; CAA33366.1; JOINED.
PIR; S00979; QFHUH.
EMBL; X15307; CAA33366.1; JOINED.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       630 EWLKPK-KREVKKGKVPEEVE 649
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        20.9%;
                                                                                          Pfam; PF00038; filament; 1.
                                                                                                                                                                                                                                                                                                                                                                                 130; Conservative
                                                                             InterPro; IPR001664; IF.
                                                                                                                                                        413
1020
132
145
244
266
                                                                                                                                                                                                      133 14
146 24
245 26
267 28
289 29
293 41
                                                                                                                                                                                                                                                                                                                                                                  Best Local Similarity
                                                               MIM; 162230;
                                                                                                                                                                                                                                                                                                                                                    Query Match
                                                                                                                                                                                                                                                                                                       SEQUENCE
                                                                                                                                                            DOMAIN
                                                                                                                                                                                          DOMAIN
                                                                                                                                                                                                         DOMAIN
                                                                                                                                                                                                                          DOMAIN
                                                                                                                                                                                                                                                                                         DOMAIN
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       658
                                                                                                                                                                                                                                                                                                                                                                                      Matches
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           Db
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          qq
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               qq
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              qq
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          δ
                                                                                                                                                                                                                                                                                                                                                                                                                                                g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                               ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           qq
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        Dp
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         δý
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      Pp
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       Qγ
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                qq
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  ٥y
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  \Omega
   ò
```

```
Detween the Swiss Institute of Bioinformatics and the EMBL outstation the European Holinformatics institute. There are no restrictions on its modified and this statement is not removed. Usage by another a license agreement is not removed. Usage by and for commercial entities requires a license agreement (See http://www.isb-sib.ch/announce/or send an email to license@isb-sib.ch).
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  116 IPVPSVVSREPTAPYHHKSTNAEEREAKEPVLKTHSKDVEREAEIGV------DDVVK 167
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               -i- DEVELOPMENTAL STAGE: UPON TERMINAL NEURAL DIFFERENTIATION, NESTIN IS DOWN-REGULATED AND REPLACED BY NEUROFILAMENTS.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           Best Local Similarity 18.8%; Pred. No. 4;
Matches 167; Conservative 102; Mismatches 287; Indels 334; Gaps
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      685 SPEEEDQEACRPLOKENQEPLGYEEAEGO-ILERLIEKESQESLRSPEEEDQEAGRSLOK 743
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               70 ------DANGHCLRE--GKEQEEGVKMAIEGLARMERRLQRATKRQKKQLEEDG--- 115
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         744 ENGEPLGYBEAEDOMLERLIEKESQESLKSPEEN-ORIGKPLERENOKSLRYLEENGETF 802
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              803 VPLESRNOR----PLRSLEVEEEEORIVKPLEKVSODSLGSLAEENVOPLRYLEEDDCIN 858
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          213 LNTYLRNAKPPIFSSRTCRMASIV-------54
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   904 VNHLEKESQE--FSRSSEEEEQVMERSLEGENHESLSSVEKEDQMVESQLEKESQDSGKS 961
                                                                                                                                                                                                                                 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          13 TPQSESSTESSTLDSSAPSPARNLRRSGRNILQPSSEKDRDHEKRSGEE---LAGRMMGK 69
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       --EPAATNIIEPEDAQDAAERGAARPPAVNSSYLPLPWKGRLGYAC
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    859 KSLLEDKTHKSLGSLEDRNGDSIIIPQESETQV---SLRPPEEEDQRI------
                                                                                                                                                                                                                                                                                                                                                                                MEDLINE=90150286; PubMed=1689217;
Lendahl U., Zimmerman L.B., McKay R.D.G.;
"CNS stem cells express a new class of intermediate filament
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           -!- SIMILARITY: BELONGS TO THE INTERMEDIATE FILAMENT FAMILY.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    3.7%; Score 128; DB 1; Length 1805;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  MW; B40EE14717E0998D CRC64;
                                                     01-MAY-1991 (Rel. 18, Created)
01-MAY-1991 (Rel. 18, Last sequence update)
20-AUG-2001 (Rel. 40, Last annotation update)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        Plam: rrvvcv.
PROSITE; PS00226; IF; 1.
Intermediate 'llament; Coiled coil; Neurone.
7 HEAD.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       COIL 1A.
LINKER 1.
COIL 1B.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           TISSUE SPECIFICITY: CNS STEM CELLS.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                LINKER 12.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         COIL 2A.
LINKER 2.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        COIL 2B.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     EMBL; M34384; AAA41685.1; -.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     198744
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           Pfam; PF00038; filament; 2.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  InterPro; IPR001664; IF.
STANDARD;
                                                                                                                                                                                                           Rattus norvegicus (Rat).
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             314
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   55
151
174
193
196
314
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          43
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               Cell 60:585-595(1990).
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   56
152
175
194
197
1805 AA;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     PIR; A34736; A34736
                                                                                                                                                                                                                                                                                                                                                             SEQUENCE FROM N.A.
                                                                                                                                                                                                                                                                                                NCBI_TaxID=10116;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             31.5
8
44
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 168 M-----
NEST_RAT
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        SEQUENCE
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 Query Match
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      protein.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          DOMAIN
DOMAIN
DOMAIN
DOMAIN
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    DOMAIN
                                                                                                                                                  NESTIN.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      QD
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       QQ
      S THE TETTE TO BRANCH OF STANK AND STANCE OF STANCE OF STANCE OF STANCE OF STANCE OF STANCE OF STANCE OF STANCE OF STANCE OF STANCE OF STANCE OF STANCE OF STANCE OF STANCE OF STANCE OF STANCE OF STANCE OF STANCE OF STANCE OF STANCE OF STANCE OF STANCE OF STANCE OF STANCE OF STANCE OF STANCE OF STANCE OF STANCE OF STANCE OF STANCE OF STANCE OF STANCE OF STANCE OF STANCE OF STANCE OF STANCE OF STANCE OF STANCE OF STANCE OF STANCE OF STANCE OF STANCE OF STANCE OF STANCE OF STANCE OF STANCE OF STANCE OF STANCE OF STANCE OF STANCE OF STANCE OF STANCE OF STANCE OF STANCE OF STANCE OF STANCE OF STANCE OF STANCE OF STANCE OF STANCE OF STANCE OF STANCE OF STANCE OF STANCE OF STANCE OF STANCE OF STANCE OF STANCE OF STANCE OF STANCE OF STANCE OF STANCE OF STANCE OF STANCE OF STANCE OF STANCE OF STANCE OF STANCE OF STANCE OF STANCE OF STANCE OF STANCE OF STANCE OF STANCE OF STANCE OF STANCE OF STANCE OF STANCE OF STANCE OF STANCE OF STANCE OF STANCE OF STANCE OF STANCE OF STANCE OF STANCE OF STANCE OF STANCE OF STANCE OF STANCE OF STANCE OF STANCE OF STANCE OF STANCE OF STANCE OF STANCE OF STANCE OF STANCE OF STANCE OF STANCE OF STANCE OF STANCE OF STANCE OF STANCE OF STANCE OF STANCE OF STANCE OF STANCE OF STANCE OF STANCE OF STANCE OF STANCE OF STANCE OF STANCE OF STANCE OF STANCE OF STANCE OF STANCE OF STANCE OF STANCE OF STANCE OF STANCE OF STANCE OF STANCE OF STANCE OF STANCE OF STANCE OF STANCE OF STANCE OF STANCE OF STANCE OF STANCE OF STANCE OF STANCE OF STANCE OF STANCE OF STANCE OF STANCE OF STANCE OF STANCE OF STANCE OF STANCE OF STANCE OF STANCE OF STANCE OF STANCE OF STANCE OF STANCE OF STANCE OF STANCE OF STANCE OF STANCE OF STANCE OF STANCE OF STANCE OF STANCE OF STANCE OF STANCE OF STANCE OF STANCE OF STANCE OF STANCE OF STANCE OF STANCE OF STANCE OF STANCE OF STANCE OF STANCE OF STANCE OF STANCE OF STANCE OF STANCE OF STANCE OF STANCE OF STANCE OF STANCE OF STANCE OF STANCE OF STANCE OF STANCE OF STANCE OF STANCE OF STANCE OF STANCE OF STANCE OF STANCE OF STANCE OF STANCE OF STANCE OF STANCE OF STANCE OF S
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            δŏ
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               q
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              ŏλ
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    q
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 δ
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   ò
```

RESULT 13

NEST_RAT

```
1022 LPKPLGNDQEIARSLGKENQESLVSL----KEKGIETVKSLETEIIEPLETAEEDLERRK 1077
                                                                                                                                                                                                           1078 SIDTQEPLWSTEVARETVEPPEDEPPGSLGSV-DENRETLTSLEKESQELSSLGKWNVET 1136
                                                                                                                                                                                                                                                                                            1137 RVEDSQQCLQVEEGLQEEQHQESLREVKQELPSSGNQQRWEDVVEGKAVGQEAPLATTGV 1196
                                                                                                                                                                                                                                                                                                                                                                                1197 GTEDKAELHLRGGGGEEEAAAEG-------ELLGDIVGEAWSLGSSEP--K 1238
                                                                                                                                                                                                                                                                                                                                                                                                                                                                 1239 EQRVPAEALD-----NLEGGALEVPVAQSMPEVTERDEDRAQAGEQDSIEVTLG 1287
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  1288 LEAARTGLELEQEVVGLEDPR----HPAREEAIPPSLGEESVKAKIAQGLEGPGKEPKEA 1343
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             1344 GALDSGILELPKTSSEALECOGHEESESMEGWEEEEASLETSDHEGSDAPQPRPPETEED 1403
                                        962 LEDESQETFGPLEKENAESLRSLAGODOEEQKLEQETQOTLRAVGNEQMAVSPPEKVDPE 1021
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                1404 EGAQAALTAPGPKLLEPCSPIPILTDAHELQPQAEGIQEAGWOPEAGSEALERVENEPEF 1463
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       NCRI_MOUSE STANDARD; PRT; 2453 AA.
060974; 060812;
20-AUG-2001 (Rel. 40, Created)
20-AUG-2001 (Rel. 40, Last sequence update)
20-AUG-2001 (Rel. 40, Last annotation update)
NUCLEAR RECEPPOR CO-REPRESSOR 1 (N-COR1) (N-COR) (RETINOID X RECEPTOR INTERACTING PROTEIN 13) (RIP13).
                                                                                                                                                                                                                                                                                                                                       385 -----MIIHMGGQFGDKAATLERFKRNYARLSQSCKNRLVLENDDVG---WTVHDLLPVCE 437
                                                                                                                                                                                                                                                                                                                                                                                                                           438 ELNIPM-VLDYHHHNICFDPAHLREGTLDI----SDPKLQER-----IANTWK 480
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              481 RKGIKQKMHYSEPCDGAVTPRDRRKHRPRVMTLPPCPPDMDLM-----IEAKDKE---- 530
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    531 ----QAVFELMRT----FKLPGFEKINDMVPYDR------DDENRPAPPVKAP--KKK 572
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     573 KGGKRKRT-------TD-----TD-----EEAAEPEEVDTAADDVKDAPE- 603
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            Seol W., Choi H.S., Moore D.D.;
"Isolation of proteins that interact specifically with the retinoid X.
receptor: two novel orphan receptors.";
                                                                                                                                                                  305 ----PFASHPVHGYKLAPFASEVLAEAGRVAAELGHRLTTHPGQFTQLGSPRKEVVES
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 Mus musculus (Mouse).
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
VCBI_TaxID=10090;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            "Ligand-independent repression by the thyroid hormone receptor mediated by a nuclear receptor co-repressor.":
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            Hoerlein A.J., Naeaer A.M., Heinzel T., Torchia J., Gloss B.,
Kurokawa R., Ryan A., Kamei Y., Soederstroem M., Glass C.K.,
Rosenfeld M.G.;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  1464 GLGEIPE------GLODW------EEGREESEADDLGE 1489
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         604 GPKEVPEEERAMGGPYNRVYWPLGCEEWLKPKKREVKKGKVPEEVEDEGE 653
                                                                                                                                                                                                                                                      359 AIRDL-----DAV---
                                                                                   ------LGLANARDIVKMLCWNEKYGIRFLR-LSSEMF-
SEQUENCE OF 1792-2453 FROM N.A. (ISOFORM LONG).
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       receptor co-repressor.";
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             SEQUENCE FROM N.A. (ISOFORMS LONG AND SHORT).
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      TISSUE=Liver;
MEDLINE=95280959; PubMed=7760852;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      MEDLINE=96008539; PubMed=7566114;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         Nature 377:397-404(1995).
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    TISSUE=Pituitary;
255 KPDKSKEP---
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 RESULT 14
NCR1_MOUSE
                                        g
                                                                                   õ
                                                                                                                       g
                                                                                                                                                                    ò
                                                                                                                                                                                                             g
                                                                                                                                                                                                                                                        ò
                                                                                                                                                                                                                                                                                            임
                                                                                                                                                                                                                                                                                                                                         ö
                                                                                                                                                                                                                                                                                                                                                                                qq
                                                                                                                                                                                                                                                                                                                                                                                                                           ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                   q
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              QQ
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                5
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             ô
```

```
This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation - the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See http://www.isb-sib.ch/announce/or send an email to license@isb-sib.ch).
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          28;
. Endocrinol. 9:72-85(1995).
FUNCTION: MEDIATES THE TRANSCRIPTIONAL REPRESSION ACTIVITY OF SOME
                                                               SUBUNIT: FORMS A LARGE COREPRESSOR COMPLEX THAT CONTAINS SINSA/B AND HISTONE DEACETYLASES HDACI AND HDAC2. THIS COMPLEX ASSOCIATES WITH THE THYROID (TR) AND THE RETINOID ACID RECEPTORS (RAR) IN THE
                                                                                                                                                                                                                                                                              DOMAIN: THE TWO INTERACTION DOMAINS (ID) CONTAIN A CONSERVED SEQUENCE REFERRED TO AS THE CORNR BOX. THIS MOTIF IS REQUIRED AND SUFFICIENT TO PERMIT BINDING TO UNLIGANDED TR AND RARS. SEQUENCES
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               (BY SIMILARITY).
                                                                                                                                                              SHORT FORM; ARE PRODUCED BY ALTERNATIVE SPLICING.
-1-TISSUE SPECIFICITY: UBIQUITOUS.
-1-DOMAIN: THE N-TERMINAL REGION CONTAINS REPRESSION FUNCTIONS THAT ARE DIVIDED INTO THREE INDEPENDANT REPRESSION DOMAINS (RDI, RD2 AND RD3). THE C-TERMINAL REGION CONTAINS THE NUCLEAR RECEPTOR-INTERACTING DOMAINS THAT ARE DIVIDED IN TWO SEPARATE INTERACTION
                                                                                                                                  SUBCELLULAR LOCATION: NUCLEAR.
ALTERNATIVE PRODUCTS: 2 ISOFORMS; A LONG FORM (SHOWN HERE) AND A
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    103 SLESKRPRLE--QVSDSHFQRI--SAAVLPLVHTLPEG---LRSSANAKKDPAFGVKHEA 155
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     57 RSGEELAGRMMGKDANGHCLREGKEQEEGVKMAIEGLARMERRLQRA-----TKRQKKQ 110
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      Indels 225; Gaps
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    3 SRKSKAAALDTPQSESSTFSSTLDSSAPSPARNLRRSGRNILQPSSEKDRD-----HEK 56
                                                                                                                                                                                                                                                                                                                                                                 -!- SIMILARITY: CONTAINS 1 SANT DOMAIN (COMPOSED OF SANT-A+SANT-B).
-!- SIMILARITY: CONTAINS 2 CORNR BOX.
-!- SIMILARITY: BELONGS TO THE N-COR NUCLEAR RECEPTOR COREPRESSORS
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         PROSITE; PS50090; MYB_3; 1,
Nuclear protein; Transcription regulation; DNA-binding; Repressor;
Colled coil; Alternative splicing.
                                   NUCLEAR RECEPTORS BY PROMOTING CHROMATIN CONDENSATION, THUS PREVENTING ACCESS OF THE BASAL TRANSCRIPTION.
                                                                                                                                                                                                                                                                                                                                   FLANKING THE CORNR BOX DETERMINE NUCLEAR HORMONE RECEPTOR
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      DB 1; Length 2453;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  52208B40382F7E6A CRC64;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                MISSING (IN SHORT ISOFORM)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             INTERACTION WITH SINJA/B COILED COIL (POTENTIAL).
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             SANT-A (POTENTIAL).
SANT-B (POTENTIAL).
COILED COIL (POTENTIAL).
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             COILED COIL (POTENTIAL)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      80; Mismatches 232;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    I -> T (IN REF. 2).
A -> P (IN REF. 2).
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                CORNR BOX OF 1D2
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                CORNR BOX OF ID1
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     3.7%; Score 127.5;
18.6%; Pred. No. 6.3
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 POLY - ALA.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               POLY-GLN.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               POLY-ALA.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                PRO-RICH
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 POLY - PRO.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  POLY-SER
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         InterPro; IPR001005; Myb_DNA_bind.
Pfam; PF00249; myb_DNA-binding; 2.
SMART; SM00395; SANT; 2.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            Alternative splicing
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  270640 MW;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          EMBL; U35312; AAB17125.1; -.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             AAC52168.1; -.
                                                                                                                                                                                                                                                                    DOMAINS (ID1 AND ID2)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 Matches 123; Conservative
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        MGD; MGI:1349717; NCOrl
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         550
616
2077
2281
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         482
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             1718
                                                                                                                     ABSENCE OF LIGAND.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             328
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              602
1047
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                1952
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  2090
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                2371
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                2453 AA;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   Best Local Similarity
                                                                                                                                                                                                                                                                                                                                                     SPECIFICITY.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              1952
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             8961
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              58
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               1044
                                                                                                                                                                                                                                                                                                                                                                                                                        FAMILY
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           DNA_BIND
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             DNA_BIND
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                SEQUENCE
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   Query Match
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                VARSPLIC
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  CONFLICT
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  CONFLICT
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        DOMAIN
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            DOMAIN
DOMAIN
DOMAIN
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            DOMAIN
DOMAIN
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            DOMAIN
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           DOMAIN
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            DOMAIN
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               DOMAIN
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             DOMAIN
   ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       ô
```

```
534 DDEEKD----DKEDSKETTKEKDRTEATAEEPEEREGV-----TPRGRKTANSQGRGKG 583
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         434 ROFMNVWTDHEKEIFKDKFIOHPKNFGLIASYLERKSVPDCVLYYYLTKKNENYK---- 488
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  497 AVTPRDRRKHRPRVMTLPPCPPDMDLMIEAKDKEQAVFELMRTFKLPGFEKINDMVPYDR 556
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      557 DDENRPAPPVKAPKKKKGGKRKRTIDEEAAEPEEVDIAADDVKDAPEGPKEVPEEERAMG 616
                                                                                                                                                                                                                                                                                                                                                                  -----NN 328
                                                                                                                                                                                                                                                                                                                                                                                                          351 PRKEVVESAIRDLEYHDELLSLLKLPEQQNRDAVMIIHMGGQFGDKAATLERFKRNYARL 410
                                                                                                                                                                                                                                                                                                                                                                                                                                                   329 PRRKAKESKTREY-YEKQFPEIRKQREQQER----FORVGORGAGLSATIARSEHEISEI 383
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            411 SQSCKNRLVLENDDVGWTVHDLLPVCEELNIPMVLDYHHHNICF------DPA---- 457
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     384 IDGLSEQ---ENNEKQMRQLSVIP-----PMMFDAEQRRVKFINMNGLMEDPMKVYKD 433
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                156 PS-SPLSGQPCGDDQNASPSKLSKEE-----LIQSMDRVDREIAKVEQQILKLKKKQQQ 208
                                                           111 LEEDGIPVPSVVSRFPTAPYHHKSTNAEEREAKEPVLKTHSKDVEREAEIGVDDVVKMEP 170
                                                                                                                                               171 AATNIIEPEDAQDAAERGAARPPAVNSSYLPLPWKGRLGYACLNTYLRNAKPPIFSSRTC 230
                                                                                                                                                                                                                                  231 RMASIVDHRHPLQFEDEPEHHLKNKPDKSKEPQDELGHKFVQELGLANARDIVKMLCWNE 290
                                                                                                                                                                                                                                                                              291 KYGIRFLRLSSEMFPFASHPVHGYKLAPFASEVLAEAGRVAAELGHRLTTHPGQFTQLGS 350
                                                                                                                                                                                          -----TNQVM 285
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              Murakumo Y., Rasio D., Roth T., Negrini M., Croce C.M., Fishel R., "Cloning and characterization of hREV3, the human homolog of S.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      Lin W., Wu X., Wang Z.; "A full-length cDNA of hREV3 is predicted to encode DNA polymerase
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          MEDLINE-98284025; PubMed-9618506; Gibbs P.E.M., McGregor W.G., Maher V.M., Nisson P., Lawrence C.W.; Gibbs P.E.M., McGregor W.G., Maher V.M., Nisson P., Lawrence C.W.; Ah human homolog of the Saccharomyces cerevisiae REV3 gene, which encodes the catalytic subunit of DNA polymerase zeta."; Proc. Natl. Acad. Sci. U.S.A. 95:6876-6880(1998).
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             Eukaryota, Metazoa, Chordata, Craniata, Vertebrata, Buteleostomi,
Mammalia; Eutheria; Primates; Catarrhini, Hominidae; Homo.
                                                                                                      209 LEEEAAKPPE----PEKPVSPPPVEQKHRSIVQIIYDENRKKAEEAHKIFEGLGPKVEL
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   DPOZ_HUMAN STANDARD; PRT; 3130 AA. 060673; 043214; 15-DEC-1998 (Rel. 37, Last sequence update) 20-405-2001 (Rel. 40, Last annotation update) DNA POLYMERASE ZETA CATALYTIC SUBUNIT (EC 2.7.7.7) (HREV3). REV3L OR POLZ OR REV3.
                                                                                                                                                                                                                                                                                                                                                                    -----VDRIE-----
···
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               zeta for damage-induced mutagenesis in humans.";
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           TISSUE-Bone marrow, and Leukocyte;
MEDLINE=99202263; PubMed=10102035;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   Mutat. Res. 433:89-98(1999).
                                                                                                                                                                                            264 PLYN--QPSDTK------
                                                                                                                                                                                                                                                                                                                                                                      321 K-----
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               Homo sapiens (Human)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            TISSUE-Fetal brain;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 SEQUENCE FROM N.A.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           SEQUENCE FROM N.A.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     SEQUENCE FROM N.A.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             NCBI_TaxID=9606;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       DPOZ_HUMAN
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     qq
                                                                                                                                                                                                                                                                              QQ
                                                                                                                                                                                                                                                                                                                                                                    q
                                                                                                                                                                                                                                                                                                                                                                                                               Qγ
                                                                                                                                                                                                                                                                                                                                                                                                                                                      qq
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 δ
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     Оþ
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       δλ
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              q
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             δ
                                                                                                                                                                                                                                         δ
                                                                                                                                                                                                                                                                                                                             οy
```

```
This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation. The European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See http://www.isb-sib.ch/announce/
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     1493 PRSLSEAISQT----KALSQCKN-----RNVSTPSA----FGEGQSGLAVLKELLQKRQQK 1540
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    MEDLINE=99126315; PubMed=9925914;
Morelli C., Mungall A.J., Negrini M., Barbanti-Brodano G., Croce C.M.;
"Alternative splicing, genomic structure, and fine chromosome localization of REV3L.";
Cytogenet. Cell Genet. 83:18-20(1998).
-:- CATALYTIC ACTIVITY: N DEOXYNUCLEOSIDE TRIPHOSPHATE =
                                                                                                                                                                                                                                                                                                                                                                                                                           -i-SIMILARITY: BELONGS TO DNA POLYMERASE TYPE-B FAMILY.
-i-CAUTION: IT IS UNCERTAIN WHETHER MET-1 OR MET-79 IS THE INITIATOR.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               70 DANGHCLREGKEQEEGVKMAIEGLARMERRLQRATKRQKKQLEEDGIPVPSVVSRFPTAP 129
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  ilarity 19.1%; Pred. No. 9.3;
Conservative 110; Mismatches 270; Indels 260; Gaps
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              14 PQSESSTFSSTLDSSAPSPARNLRRSGRNTLQPSSEKDRDHEKRSG----EELAGRMMGK 69
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     Transferase; DNA-directed DNA polymerase; DNA replication; DNA-binding; DNA repair; Nuclear protein; Zinc-finger; Polymorphism. ZN_FING 3042 3057 C4-TYPE (POTENTIAL).
                                                                                                    Rasio D., Murakumo Y., Negrini M., Croce C.M., Fishel R.;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    3.6%; Score 127; DB 1; Length 3130;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            /FTId=VAR_008519.
E -> Q (IN REF. 4 AND 5).
Y -> C (IN REF. 4 AND 5).
AW: 1C0700900F10BB14 CRC64;
                                                                                                                                   Submitted (NOV-1997) to the EMBL/GenBank/DDBJ databases.
                        Submitted (JUN-1999) to the EMBL/GenBank/DDBJ databases
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               C4-TYPE (POTENTIAL).
C4-TYPE (POTENTIAL).
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           /FTId-VAR_008516.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      S -> T.
/FTId=VAR_008517.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        /FTIG=VAR_008518.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 or send an email to license@isb-sib.ch).
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   PROSITE; PS00116; DNA_POLYMERASE_B; 1.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  IPR002064; DNA_pol_B.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     EMBL, AF058701, AAC24357.1; -. EMBL, AF071798; AAC24009.1; -. EMBL, AF157476; AAD40184.1; -. EMBL, AF035537; AAC28460.1; -. EMBL, AF078695; AAC28460.1; -.
                                                                                  SEQUENCE OF 79-3130 FROM N.A.
                                                                                                                                                                                                      SECUENCE OF 79-3130 FROM N.A.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              Pfam; PF00136; DNA_pol_B; 2. PRINTS; PR00106; DNAPOLB.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      1156 1156
3130 AA; 352782
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            19.18;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              237
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 3104
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      389
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                1540
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       2607
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               231
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           POLBC;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     Best Local Similarity
cerevisiae REV3.";
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                1540
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       2607
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       3086
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      389
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  231
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        SM00486;
                                                                                                                                                                                                                                       TISSUE=Testis;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       602776,
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 Matches 151;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        InterPro;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          Query Match
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 SEQUENCE
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    CONFLICT
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            CONFLICT
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          VARIANT
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       ZN_FING
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  VARIANT
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            VARIANT
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                VARIANT
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              SMART;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             MIM;
      S THE TETTE TETTE SOLUTION OF THE SET TO SET TO SET TO SET TO SET TO SET TO SET TO SET TO SET TO SET TO SET TO SET TO SET TO SET TO SET TO SET TO SET TO SET TO SET TO SET TO SET TO SET TO SET TO SET TO SET TO SET TO SET TO SET TO SET TO SET TO SET TO SET TO SET TO SET TO SET TO SET TO SET TO SET TO SET TO SET TO SET TO SET TO SET TO SET TO SET TO SET TO SET TO SET TO SET TO SET TO SET TO SET TO SET TO SET TO SET TO SET TO SET TO SET TO SET TO SET TO SET TO SET TO SET TO SET TO SET TO SET TO SET TO SET TO SET TO SET TO SET TO SET TO SET TO SET TO SET TO SET TO SET TO SET TO SET TO SET TO SET TO SET TO SET TO SET TO SET TO SET TO SET TO SET TO SET TO SET TO SET TO SET TO SET TO SET TO SET TO SET TO SET TO SET TO SET TO SET TO SET TO SET TO SET TO SET TO SET TO SET TO SET TO SET TO SET TO SET TO SET TO SET TO SET TO SET TO SET TO SET TO SET TO SET TO SET TO SET TO SET TO SET TO SET TO SET TO SET TO SET TO SET TO SET TO SET TO SET TO SET TO SET TO SET TO SET TO SET TO SET TO SET TO SET TO SET TO SET TO SET TO SET TO SET TO SET TO SET TO SET TO SET TO SET TO SET TO SET TO SET TO SET TO SET TO SET TO SET TO SET TO SET TO SET TO SET TO SET TO SET TO SET TO SET TO SET TO SET TO SET TO SET TO SET TO SET TO SET TO SET TO SET TO SET TO SET TO SET TO SET TO SET TO SET TO SET TO SET TO SET TO SET TO SET TO SET TO SET TO SET TO SET TO SET TO SET TO SET TO SET TO SET TO SET TO SET TO SET TO SET TO SET TO SET TO SET TO SET TO SET TO SET TO SET TO SET TO SET TO SET TO SET TO SET TO SET TO SET TO SET TO SET TO SET TO SET TO SET TO SET TO SET TO SET TO SET TO SET TO SET TO SET TO SET TO SET TO SET TO SET TO SET TO SET TO SET TO SET TO SET TO SET TO SET TO SET TO SET TO SET TO SET TO SET TO SET TO SET TO SET TO SET TO SET TO SET TO SET TO SET TO SET TO SET TO SET TO SET TO SET TO SET TO SET TO SET TO SET TO SET TO SET TO SET TO SET TO SET TO SET TO SET TO SET TO SET TO SET TO SET TO SET TO SET TO SET TO SET TO SET TO SET TO SET TO SET TO SET TO SET TO SET TO SET TO SET TO SET TO SET TO SET TO SET TO SET TO SET TO SET TO SET TO SET
```

YHHKSTNAEEREAKEPVLKTHSKDVEREAEIGVDDVVKMEPAATN 174 	IIEPEDAQDAAERGAARPPAVNSSYLPLPWKGRLGYACLNTYLRNAKPPIF 225	SSRTCRMASIVDHRHPLQFEDEPEHHL	257 DKSKEPQDELGHKFVQELGLANARDIVKMLCWNEKYGIRFLRLSSEMFPFASHPVHGYKL 316 1	317 APFASEVLAEAGRVAAELGHRLTTHPGQFTQ-LGSPRKEVVESAIRDLEYHDELLSLLKL 375 	376 PEQQNRDAVMIIHMGGQFGDKAATLERFKRNYARLSQSCKNRLVLENDDVGWTVHDLLPV 435 1	436 CEELNIPMYLDY 477 136 CEELNIPMYLDY 477 11: : : : : :	478TWKR	VMTLP-PCPPDMDLMIBAKDKEQAVFELMRTFKLPGFEKINDMVPYDRDDENRPAPP 565 :::	VK	601APEGPKEVPEEERAMGGPYNRVYWPLGCEEWLKFKKRE 638 	
			DKSKEPQDELGHKFVQELGLANARDI 	APFASEVLAEAGRVAAELGHRLTTHP 	PEQQNRDAVMITHMGGQFGDKAATLE: 	CEELNIPMVLDYHHH : : : MSPPSREEI-MATLLDHDLSETIYQE	EGDFSLEGLRLWKTAFSAMTQNPRPG		VKAI PKMDVSPCILPTTAHTKEDVDNSQIA	DDDNYXISYSSPDSPVIPPWQQPISP	VKKGKVPEEVE 649 :: IQKSPCSEPQE 2178
130	175 1639	226 1687	257 1745	317	376	436	478	510	566	601	639
Oy Dp	Qy Dp	Qy Db	Qy Dp	Qy Db	Qy Db	oy Op	Oy Db	Oy Dp	Oy Dp	Qy	Oy Dp

Search completed: January 15, 2002, 14:07:17 Job time: 946 sec

	·		
		•	
			•
			•

```
(without alignments)
255.755 Million cell updates/sec
                                                                                                                                                                                                                                                                                               1 MPSRKSKAAALDTPQSESST.....REVKKGKVPEEVEDEGEFDG 656
                                                                                                                                               January 15, 2002, 13:51:04 ; Search time 57.72 Seconds
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         Issued_Patents_AA:*
|: /cgn2_6/ptodata/2/iaa/5A_COMB.pep:*
|: /cgn2_6/ptodata/2/iaa/5B_COMB.pep:*
|: /cgn2_6/ptodata/2/iaa/6A_COMB.pep:*
|: /cgn2_6/ptodata/2/iaa/6B_COMB.pep:*
|: /cgn2_6/ptodata/2/iaa/PCTUS_COMB.pep:*
|: /cgn2_6/ptodata/2/iaa/PCTUS_COMB.pep:*
|: /cgn2_6/ptodata/2/iaa/PCTUS_COMB.pep:*
GenCore version 4.5
Copyright (c) 1993 - 2000 Compugen Ltd.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                          Total number of hits satisfying chosen parameters:
                                                                                                                                                                                                                                                                                                                                                                                                                               212252 seqs, 22503292 residues
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries
                                                                                               OM protein - protein search, using sw model
                                                                                                                                                                                                                                                                                                                                                                            Gapop 10.0 , Gapext 0.5
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          Minimum DB seq length: 0
Maximum DB seq length: 200000000
                                                                                                                                                                                                                                                  US-09-724-296-36
3482
                                                                                                                                                                                                                                                                                                                                                    BLOSUM62
                                                                                                                                                                                                                                                                          Perfect score:
                                                                                                                                                                                                                                                                                                                                                    Scoring table:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            Database :
                                                                                                                                                                                                                                                                                                     Sequence:
                                                                                                                                                                                                                                                                                                                                                                                                                            Searched:
                                                                                                                                               Run on:
```

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

D	Description	Sequence 8. Appli		, 1	Sequence 3, Appli	ì v	, c	Ì		16.	2 4	, ,	٦.	11,	, , ,	ìσ	, <	Sequence 4, appli	ic	٠ ر	, 5	Ì	, v	, o	Sequence 0, appli	,	Sequence 3, Appli	· 🗂
SUMMARIES		US-08-923-992A-8	3-9	-08-769-309A	US-08-994-570-5	US-08-923-992A-6	US-08-923-992A-2	US-08-923-9928-4	PCT-US93-07261-11	PCT-US93-07261-16	US-08-188-582-14	US-08-646-715-14	US-08-188-582-11	US-08-646-715-11	US-08-475-844-9	PCT-US95-08429-9	3-913-4	- 1	US-08-719-822B-2	US-09-092-458-2	US-09-315-793-52	US-08-675-816-6	US-08-285-440-6	0-349-	-09-085-199R	-08-766	-09-413-814-	-09-018-628-1
DB	1	4	-	-	m	4	4	4	2	S		, ,	-		~	Ŋ	П	٦	~	4	4	4	Н		4	7	4	7
Length		1098	1805	1780	1780	1128	1164	1104	1588	1663	1872	1872	1893	1893	727	727	1618	1018	1018	1018	1093	542	558	558	1068	251	1213	410
% Query Match		3.8		٠.	3.5	3.4	3.4	3.4	3.4	3.4	3.4	3.4	3.4	3.4	3,3	3.3	3.3	3.3	3.3	3.3	3.3	3.2	3.2	3.2	3.2	3.2	3.1	3.1
Score		131.5	128	121	121	119.5	118.5	•		117.5	117.5		117.5	117.5	115	115	114.5	113.5	113.5	113.5	113.5	113	112	112	112	110.5	109.5	109
Result No.		-	2	m	4	2	9	7	80	6	10	11	12	13	14	15	16	17	18	19	20	21	22	23	24	25	26	27

Sequence 14, Appl		Sequence 16, Appl	Sequence 18, Appl	Sequence 18. April			Sequence 1. Appli	4			. 0	Segmence 48 Appl	, ,	Secuence 2 Appli	1 4	, ,	Sequence 477, Appl Sequence 5, Appli	
US-09-273-378-14	US-09-018-628-16	US-09-2/3-3/8-I6	US-09-018-628-18	US-09-273-378-18	US-09-018-635-27	US-09-120-663-2	US-08-766-738-1	US-08-533-306A-4	US-08-742-923A-4	US-08-962-284-2	US-08-455-543A-48	US-08-223-305C-48	US-08-713-118-2	US-09-452-007-2	US-08-455-543A-47	US-08-223-305C-47	US-08-285-440-5	
m	71	ς,	~	m	4	7	2	~	7	7	~	7	~	4	-	7	Н	
410	φ	400	200	200	200	1.838	251	885	885	1788	2237	2237	2337	2337	2339	2339	532	
3.1	7.0	7.0	3.1	3.1	3.1	3.1	3.1	3.1	3.1	3.1	3.1	3.1	3.1	3.1	3.1	3.1	3.0	
109	109	103	7.03	109	109	108	107.5	107	107	107	107	107	107	107	107	107	106	
28	V C	0 0	31	32	33	34	35	36	37	38	39	40	41	42	43	44	45	

ALIGNMENTS

```
APPLICANT: Tai, Joseph Y.
APPLICANT: Blake, Milan S.
TITLE OF INVENTION: No. 6280738-IGA FC Binding Forms of the Group B
TITLE OF INVENTION: Streptococcal Beta Antigens
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          Gaps
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 Ouery Match
Best Local Similarity 18.4%; Pred. No. 0.0042;
Matches 121; Conservative 94; Mismatches 259; Indels 183;
                                                                                                                                                                                                                                             ADDRESSEE: Sterne, Kessler, Goldstein & Fox P.L.L.C.
STREET: 1100 New York Avenue, N.W., Suite 600
CITY: Washington
                                                                                                                                                                                                                                                                                                                                                                                                                                                                      SOFTWARE: Patentin Release #1.0, Version #1.30 CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/923,992A FILING DATE: 05-SEP-1997 CLASSIFICATION: 536 PRIOR APPLICATION DATA:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         REFERENCE/DOCKET NUMBER: 1438.0140001/RWE TELECOMMUNICATION INFORMATION:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          APPLICATION NUMBER: US 60/024,707 FILING DATE: 06-SEP-1996 ATTORNEY/AGENT INFORMATION:
                                                                                                                                                                                                                                                                                                                                                                                                                              COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
                                            ; Sequence 8, Application US/08923992A
; Patent No. 6280738
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  /AGENT INFORMAL ESMOND, Robert W. MINMBER: 32,893
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   TELEPHONE: (202) 371-2600
TELEFAX: (202) 371-2600
INFORMATION FOR SEQ ID NO: 8:
SEQUENCE CHARACTERISTICS:
                                                                                                                                                                                                                                                                                                                                                                                                           Floppy disk
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             1098 amino acids
                                                                                                                                                                                                   NUMBER OF SEQUENCES: 3
CORRESPONDENCE ADDRESS:
                                                                                                                                                                                                                                                                                                                                                                                    COMPUTER READABLE FORM:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             MOLECULE TYPE: protein
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    REGISTRATION NUMBER:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    amino acid
                                                                                    GENERAL INFORMATION:
                                                                                                                                                                                                                                                                                                                                                                                                         MEDIUM TYPE:
                                                                                                                                                                                                                                                                                                                  D.C.
                                                                                                                                                                                                                                                                                                                                                          20005
RESULT 1
US-08-923-992A-8
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   TOPOLOGY:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    US-08-923-992A-8
                                                                                                                                                                                                                                                                                                                                       COUNTRY:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               LENGIH:
                                                                                                                                                                                                                                                                                                                  STATE:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            NAME:
```

13 TPQSESSTFSSTLDSSAPSPARNLRRSGRNILQPSSEKDRDHEKRSGEELA--GRMMGKD 70

δλ

us-09-724-296-36.rai

Db 31	
Qy 71	ANGHCLREGKEQEEGVKMAIEGLARMERRLQRATK
Db 91	VDNTILSHEQKNEFKTKIDETNDSDALLELENQFNETNRLLHIKQHEEVEKDKKAKQ
Oy 109 Db 149	KOLEEDGIPVPSVVSRFPTAPYHHKSTNAE-EREAKEPVLKTHSKDVEREAEIGVDDVVK 167 :
Qy 168	MEPAATNIIEPEDAQDAAERGAARPPAVNSSYLPLPWKGRLGYACLNTYLRNAKPPIFSS 227
Db 198	QPDKKEDAEVKVREELGKLFSS 219
Oy 228	RTCRMASIVDHRHPLOFEDEPEHHLKNKPDKSKEPQDELGHKFVOELGLA 27
Db 220	TKAGLDOELQEHVKKETSSEENTOKVDEHYANSLONLAQKSLEELDKA 267
Qy 278	NARDIVKMLCWNEKYGIRFLRLSSEMFPFASHPVHGYKLAPFASEVLAEAGRVAAELGHR 33
Db 268	TINEQATQVKNQFLENAQKLKEIQPLIKETNVKLYKAMSESLEQVEKELKH- 3
Qy 338	ב
Db 319	NSEANLEDLVAKSKEIVREYEGKLNQSKNLPELKQLEEEAHSKLKQVVEDFRKKF 373
Qy 398	ATLERESQSCKNRLVLENDDVGWTVHDLLPVCEELNIP 442
Db 374	TISEQVTPKKRLKRDLAANENNQQKIELTVSPENITVYEGED
Qy 443	MVLDYHHHNICFDPAHLREGTLDISDPKLQERI
Db 421	TAKSDSKTTLDFSDLLTKYNPSVSDRISTNYKTNTDNHKIA 461
Qy 497	7 AVTPRDRRKHRPRVWTLPPCPPDMDLMIEAKDKEQAVFELMRTFKLPGFEKINDWVP 553
.Db 462	EITIKNLKLNESQTV
Qy 554	4 YDRDDENRPAPPVKAPKKKGGKRKRTTDEEAAEPEEVDTAADDVKDAPEGPKEVPE 610
Db 509	9 EQKDSKTEEKVP-QEPKSNDKNQLQELIKSAQQELEKLEKAIKELMEQPEIPSN-PE 563
RESULT 2 US-07-853- ; Sequence	2 3-913-2 ce 2. Application US/07853913
GE	Z
; APPI	Lendahl, Urban
TITLE	LE OF INVENTION: Nestin Expression As An Indicator of LE OF INVENTION: Neuroepithelial Tumors
NUMBER	OF SEQUENCES: 4
	ADDRESS: Hamilton, B
	STREEF: Two Militia Drive CITY: Lexington
	TATE: Massachusetts ONNTRY: U.S.A.
	02173
WOO	COMPUTER READABLE FORM: MEDIUM TYPE: Floppy disk
Ō	IBM PC CON
 O Ø	ARE: Patentin Release
., cur	RKENI AFFLICATION DATA: APPLICATION NUMBER: US/07/853,913
т. С	31
PRIOR	APPLICATION DATA:
∢ 1.	APPLICATION NUMBER: US U//66U,412 FTLING DATE: 22-FEB-1991

```
41;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             1078 SIDTQEPLMSTEVARETVEPPEDEPPGSLGSV-DENRETLTSLEKESQELSSLGKWNVET 1136
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     1137 RVEDSQQCLQVEECLQEEQHQESLREVKQELPSSGNQQRWEDVVEGKAVGQEAPLATTGV 1196
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           1197 CTEDKAELHLRGQGGEEEAAAEG--------ELLQDIVGEAWSLGSSEP--K 1238
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          1239 EQRVPAEALD------NLEGGALEVPVAOSMPEVTERDEDRAQAGEQDSIEVTLG 1287
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           1022 LPKPLGNDQEIARSLGKENQESLVSL----KEKGIETVKSLETEIIEPLETAEEDLERRK 1077
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        962 LEDESQETFGPLEKENAESLRSLAGODQEEQKLEGETQQTLRAVGNEQMAVSPPEKVDPE 1021
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               305 -----PEASHPVHGYKLAPFASEVLAEAGRVAAELGHRLTTHPGQFTQLGSPRKEVVES 358
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       385 -----MIIHMGGQFGDKAATLERFKRNYARLSQSCKNRLVLENDDVG--WTVHDLLPVCE 437
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            438 ELNIPM-VLDYHHHNICFDPAHLREGTLDI----SDPKLQER-----IANTWK 480
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 481 RKGIKQKMHYSEPCDGAVTPRDRRKHRPRVMTLPPCPPDMDLM-----IEAKDKE----530
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            Query Match
Best Local Similarity 18 8%; Pred. No. 0.018;
Matches 167; Conservative 102; Mismatches 287; Indels 334; Gaps
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       274 -------LGLANARDIVKMLCWNEKYGIRFLR-LSSEMF------ 304
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   685 SPEEEDQEACRPLQKENDEPLGYEEAEGQ-ILERLIEKESQESLRSPEEEDQEAGRSLQK 743
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          70 -----DANGHCLRE--GKEQEEGVKMAIEGLARMERRLQRATKRQKKQLEEDG--- 115
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        744 ENQEPLGYEEAEDQMLERLIEKESQESLKSPEEN-QRIGKPLERENQKSLRYLEENQETF 802
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            116 IPVPSVVSREPTAPYHHKSTNAEEREAKEPVLKTHSKDVEREAEIGV------DDVVK 167
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               168 M------EPAATNIIEPEDAQDAAERGAARPPAVNSSYLPLPWKGRLGYAC 212
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             859 KSLLEDKTHKSLGSLEDRNGDSIIIPQESETOV---SLRPPEREDQRI------- 903
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  213 LNTYLRNAKPPIESSRTCRMASIV------DHRHPLQFEDEPEHHLKN 254
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   904 VNHLEKESQE-~FSRSSEEEEQVMERSLEGENHESLSSVEKEDQMVESQLEKESQDSGKS 961
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                803 VPLESRNOR----PLRSLEVEEEEQRIVKPLEKVSODSLGSLAEENVOPLRYLEEDDCIN 858
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    13 TPQSESSTESSTLDSSAPSPARNLRRSGRNILQPSSEKDRDHEKRSGEE---LAGRMMGK 69
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       359 AIRDL------EYHDELLSLLK--LPEQQNR----DAV-----
                                                                                                                                                                                                                                           RECISTRATION NUMBER: 32,227
REPERENCE/DOCKET NUMBER: MIT-4641AAAA
TELEPOOMMUNICATION INPORMATION:
TELEPHONE: 617-861-6240
TELEPAX: 617-861-6240
INFORMATION FOR SEQ ID NO: 2:
SEQUENCE CRARACTERISTICS:
                           APPLICATION DATE: 25-001.100.
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 07/201,762
                   APPLICATION NUMBER: US 07/603,803 FILING DATE: 25-0CT-1990
                                                                                                                                                                            APPLICATION NUMBER: US 07/180,548
                                                                                                                                                                                                  ATTORNEY/AGENT INFORMATION:
NAME: Granahan, Patricia
REGISTRATION NUMBER: 32,2
                                                                                                                                                                                                                                                                                                                                                                                                                                             1805 amino acids
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       MOLECULE TYPE: protein
PRIOR APPLICATION DATA:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                   AMINO ACID
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  Linear
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            TOPOLOGY:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                US-07-853-913-2
                                                                                                                                                                                                                                                                                                                                                                                                                                             LENGTH:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      Dp
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           QQ
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               ŏλ
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   Db
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       ŏ
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          δy
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               ŏ
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      qq
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      qq
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               Öλ
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              QQ
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          Q_{\lambda}
```

LEARTGLELEQEVGLEDPRHFAREEAIPPSLGEESVKAKIAQGLEGPGKEPKEA 1343QAVFELMRTFKLPGFEKINDMVPYDR	98769309A 1 J., 1 J., 1 Ssa M. Cein Binding Domains of Gravin O'Toole, Gerstein, Murray & Borun	merica -bos #1.0, Version #1.30 69,309A E.	tch al Similarity 18.3%; Score 121; DB 1; Length 1780; 134; Conservative 95; Mismatches 283; Indels 220; Gaps 30; RKSKAAALDTPQSESSTESSTLDSSAPSPARNLRRSGRNILQPSSEKDRDHEKR 57
	3 fog-309A-5 ence 5, Application US/08769309 nt No. 5741890 ERAL INFORMATION: PPLICANT: Scott, John D., PPLICANT: Nauert, Brian J., ITLE OF INVENTION: Protein Bin UMBER OF SEQUENCES: 24 ORRESPONDENCE ADDRESS: ADDRESSEE: Marshall, O'Toole, STREET: 6300 Sears Tower/233	Chicago Illinois Conceded States of Americo GoGO-6402 READABLE FORM: R: IBM PC compatible R: SYSTEM: PC-DOS/MS-DOS E: PatentIn Release #1.0, PPLICATION DATA: CATION NUMBER: US/08/769,309 DATE: CATION NUMBER: 35,302 AGENT INFORMATION: RT/DOCKET NUMBER: 35,302 ATON NU	Query Match Best Local Similarity 18.3%; Pred. N Matches 134; Conservative 95; Mism 4 RKSKAAALDTPQSESSTFSSTLDSS:
09 09 09 09	RESULT US-08-1 Seque Result Re	0.5-0	Ov Be Oy Oy Oy

```
171 AATNIIEPEDAQDAAERGAARPPAVNSSY--LPLPWKGRLGYACLNTYLRNAKPPIFSSR 228
                                                                                1: : | | : : | 345 ASEKLTASEQAHPQEPAESAHEPRLSAEXVELPSEEGVS-----GSQGPSEEKP 395
                                                                                                                                                 229 TCRMASIVDHRHPL-QFEDEPEHHLKNKPDKSKE------PQDELGHKFVQELG 275
                                                                                                                                                                                       276 LANARDIVKMLCWNEKYGIRFLRLSSEMFPFASHPVHGYKLAPFASEVLAEAGRVAAELG 335
                                                                                                                                                                                                                                                                                                                                                            336 HRLTTHPGQFTQLGSPRKEVVESAIRDLEYHDELLSLLKLPEQQNRDAVMIIHMGGQFGD 395
                                                                                                                                                                                                                                                                                                                                                                                            500 -EMLSSQERWKVQGSPLKKLFTST------GLKKLSGKRQKG-----KRGG--GD 540
                                                                                                                                                                                                                                                                                                           456 AEPAKELVKL-------KETCVSGEDPTQGADLSPDEKVLSKPPEGVVSEV- 499
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                541 E------BSP-----BSGE-----HTQVPA----DSP-----D 555
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              555 DRDDENRPAPPVKAPKK------KKGGKRKRTTDEE-----AAEPEEV 591
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              656 MKGSVEEPKP--EEPKRKVDTSVSWEALICVGSSKKRARRRSSSDEEGGPKAMGGDHQKA 713
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  592 DTAADDVKDAPEG-----PKEVPEEERAMGGPYNRVYWPLGCEEWLKPKK----RE 638
285 SPVTSETGSTFKKFFTQGWAGWRKKTSFRKPKEDEVEASEKKKEQEPEKVDTEEDGKAEV 344
                                                                                                                                                                                                                                                                                                                                                                                                                                                                  396 KAATLERFKRNYARLSQSCKNRLVLENDDVGWTVHDLLPVCEELNIPMVLDYHHHNICFD 455
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       556 SOEEQKGESSASSPEEPEEI--TCLEKGLAEVQQDGEAEEGATSDGEKKREGVTPWASFK 613
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          498 --VTPRDRRKHRPRVMTLPPCPPDMDLMIEAKDKEQAVFELMRTFKLPGFEK-INDMVPY 554
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             614 KMVTPK-KRVRRP------SESDKEDEL-DKVKSATLSSTESTASEMQEE 655
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   714 DEAGKDKETGTDGILAGSQEHDPGQGSSSPEQAGSPTEGE----GVSTWESFKRLVTPRK 769
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       E: Marshall, O'Toole, Gerstein, Murray & Borun 6300 Sears Tower/233 South Wacker Drive
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        GENERAL INFORMATION:
APPLICANT: Scott, John D.,
APPLICANT: Nauert, Brian J.,
APPLICANT: Klauck, Theresa M.
TITLE OF INVENTION: Protein Binding Domains of Gravin NUMBER OF SEQUENCES: 24
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       NAME: NO. 6090929and, Greta E. REGISTRATION NUMBER: 35,302
REFERENCE/DOCKET NUMBER: 27866/33451
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              COUNTRY: United States of America
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      APPLICATION NUMBER: US/08/994,570
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          ; Sequence 5, Application US/08994570 ; Patent No. 6090929
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  COMPUTER: IBM PC compatible
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       ATTORNEY/AGENT INFORMATION:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               CORRESPONDENCE ADDRESS: ADDRESS:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        639 VKKGKVPEEVED 650
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        | |: |: ||
770 KSKSKLEEKSED 781
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        CITY: Chicago
STATE: Illinois
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    60606-6402
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         CLASSIFICATION:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   FILING DATE:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               US-08-994-570-5
                                                                                                                                                                                                                                                                                                        gg
                                                                                                                                                                                                                                                                                                                                                                                                            qq
                                                                                                                                                                                                                                                                                                                                                                                                                                                               ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                Dp
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    Dp
                                                                                                                                                                                                                                                                                                                                                            δy
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      δ
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          οy
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          qq
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              δy
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              QQ
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           ó
```

```
30;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             500 -EMLSSQERMKVQGSPLKKLFTST------GLKKLSGKKQKG----KRGG--GD 540
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    396 KAATLERFKRNYARLSQSCKNRLVLENDDVGWTVHDLLPVCEELNIPMVLDYHHHNICFD 455
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      SQEEQKGESSASSPEEPEEI - - TCLEKGLAEVQQDGEAEEGATSDGEKKREGVTPWASFK 613
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                614 KMVTPK-KRVRRP-----SESDKEDEL-DKVKSATLSSTESTASEMQEE 655
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      555 DRDDENRPAPPVKAPKK -------KKGGKRKRTTDEE-----AAEPEEV 591
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          656 MKGSVEEPKP--EEPKRKVDTSVSWEALICVGSSKKRARRRSSSDEEGGPKAMGGDHQKA 713
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    592 DTAADDVKDAPEG-----PKEVPEEERAMGGPYNRVYWPLGCEEWLKPKK----RE 638
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        714 DEAGKDKETGTDGILAGSQEHDPGQGSSSPEQAGSPTEGE----GVSTWESFKRLVTPRK 769
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                276 LANARDIVKMLCWNEKYGIRFLRLSSEMFPFASHPVHGYKLAPFASEVLAEAGRVAAELG 335
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  456 AEPAKELVKL-------KETCVSGEDPTQGADLSPDEKVLSKPPEGVVSEV- 499
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            336 HRLTTHPGQFTQLGSPRKEVVESAIRDLEYHDELLSLLKLPEQQNRDAVMIIHMGGQFGD 395
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     541 E-----DSP----DSP----D 555
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              498 --VTPRDRRKHRPRVMTLPPCPPDMDLMIEAKDKEQAVFELMRTFKLPGFEK-INDMVPY 554
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  185 KKDKTEKPDTVQLLTVKKDEGEGAAGAGDHQDPSLGAGEAASKESEPKQSTEKPEETLKR 244
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             116 IPVP----SVVSRFPTAPYHHKSTNAEEREAKEPVLKTHSKDVEREAE-IGVDDVVKMEP 170
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   285 SPVTSETGSTFKKFFTQGWAGWRKKTSFRKPKEDEVEASEKKKEQEPEKVDTEEDGKAEV 344
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             171 AATNIIEPEDAQDAAERGAARPPAVNSSY--LPLPWKGRLGYACLNTYLRNAKPPIFSSR 228
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        345 ASEKLTASEQAHPQEPAESAHEPRLSAEYEKVELPSEEQVS------GSQGPSEEKP 395
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  229 TCRMASIVDHRHPL-QFEDEPEHHLKNKPDKSKE------PQDELGHKFVQELG 275
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            396 APLATEVFDEKIEVHQEEVVAEVHVSTVEERTEEQKTEVEETAGSVPAEELVGMDAEPQE 455
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         58 --SGEELAGRAMGKDANGHCLREGKEQEEGVKMAIEGLARMERRLQRATKRQKKQLEEDG 115
                                                                                                                                                                                                                                                                                                                                                                                                          Gaps
                                                                                                                                                                                                                                                                                                                                                                                                                                                           4 RKSKAAALDTPQ-----SESSTFSSTLDSSAPSPARNLRRSGRNILQPSSEKDRDHEKR 57
                                                                                                                                                                                                                                                                                                                                                   Query Match 3.5%; Score 121; DB 3; Length 1780; Best Local Similarity 18.3%; Pred. No. 0.079; Matches 134; Conservative 95; Mismatches 283; Indels 220;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  456 PAHLREGTLDISDPKLQERIANTWKRKGIKQKMHYSEPCDGA-
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            Sequence 6, Application US/08923992A
TELECOMMUNICATION INFORMATION:
                                           TELERA: 312-474-0448
TELEX: 25-3856
INFORMATION FOR SEQ ID NO: 5:
SEQUENCE CHARACTERISTICS:
LENGTH: 1780 amino acids
                          312-474-6300
                                                                                                                                                                                                                                       ; MOLECULE TYPE: protein US-08-994-570-5
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  639 VKKGKVPEEVED 650
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      770 KSKSKLEEKSED 781
                                                                                                                                                                                   amino acid
                                                                                                                                                                                                                      linear
                          TELEPHONE:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     RESULT 5
US-08-923-992A-6
                                                                                                                                                                                                                   TOPOLOGY:
                                                                                                                                                                                         TYPE:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             qq
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       ōλ
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               δ
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      qq
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                δŏ
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  q
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            δλ
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             q
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    ōλ
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        q
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  οy
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         ŏ
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          qq
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  Ω
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             QΥ
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              οp
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        δ
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         qq
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                οy
```

```
75 LOOWKNNLKNDVDNTIL--SHEOKNEFYTKIDETNDSDALLELENOFNETNRLLHIKOHE 132
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       322 EANLEDLVAKSKEIVREYEGKLNQSKNLPELKQLEEEAHSKLKQVVE---DFRKKFKTSE 378
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           67 ------MCKDANGHCLREGKEQEEGVKMAI-----EGLARMERRLQRATK----RQKK 109
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          110 QLEEDGIPVPSVVSRFPTAPYHHKSTNAEEREAKEPVLKTHSKDVEREAEIGVDDVVKM- 168
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             133 EVEKD------KKAKQQKTLKQSDTKVDLSNIDKELNHQKSPVEKMA 173
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      169 EP-AATN-----IIEPEDAQDAAERGAARPPAVNSSYLPLPWKGRLGYACLNTYLRNAK 221
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  174 EPKGTTNEDKDSMLKKIEDIRKQAQQADKKEDA----------EVKVREEL 214
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        222 PPIESSRTCRMASIVDHRHPLQFEDEPEHHLKNKPDKSKEPQDELGHKFVQELGLANARD 281
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              215 GKLFSSTKAGLDQEI-HEH-VKKETSSEENTQKVDEHYANSLONLAQKSLEELDKATTNE 272
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    282 IVKMLCWNEKYGIRFLRLSSEMFPFASHPVHGYKLAPFASEVLAEAGRVAAELGHRLTTH 341
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   342 PGQFTQLGSPRKEVVESAIRDLEYHDELLSLLKLPEQQNRDAVMIIHMGGQFGDKAATLE 401
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             273 QATQV---KNQFLENAQKLKEMQPLIKET--NVKLYKAMSESLEQ---VEKELKH---NS 321
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         13 TPQSESSTFSSTLDSSAPSPARNLRRSGRNILQPSSEKDRDHEKRSGEELA---GRM--- 66
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  26 TDQCNNSSSSELETTRMEIPTTDIKKA----VEPV------EKTAGETSATDTCKREKO 74
                                                APPLICANT: Tai, Joseph Y.
APPLICANT: Blake, Milan S.
TITLE OF INVENTION: No. 6280738-JgA FC Binding Forms of the Group
TITLE OF INVENTION: Streptococcal Beta Antigens
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   Mismatches 266; Indels 169;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              Ouery Match
Best Local Similarity 18.8%; Score 119.5; DB 4; Length 1128;
Best Local Similarity 18.8%; Pred. No. 0.054;
Matches 123; Conservative 95; Mismatches 266; Indels 169.
                                                                                                                                                                                                                     ADDRESSPE: Sterne, Kessler, Goldstein & Fox P.L.L.C. STREET: 1100 New York Avenue, N.W., Suite 600
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         SOFTWARE: Patentin Release #1.0, Version #1.30 CURRENT APPLICATION DATA:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        1438.0140001/RWE
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  US/08/923,992A
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    PRIORAPPLICATION DATA:
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 60/024,707
FILING DATE: 06-SEP-1996
ATTORNEY/AGENT INFORMATION:
                                                                                                                                                                                                                                                                                                                                                                                                                                                   IBM PC compatible
SYSTEM: PC-DOS/MS-DOS
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             32,893
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     REFERENCE/DOCKET NUMBER: 14
TELECOMMUNICATION INFORMATION:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    (202) 371-2600
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         TELEFAX: (202) 371-2540
INFORMATION FOR SEQ ID NO: 6:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                1128 amino acids
                                                                                                                                                                                                                                                                                                                                                                                                                           Floppy disk
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          05-SEP-1997
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     Esmond, Robert W.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     CHARACTERISTICS:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 MOLECULE TYPE: protein
                                                                                                                                                                                                                                                                                                                                                                                                COMPUTER READABLE FORM: MEDIUM TYPE: Floppy
                                                                                                                                                                                             CORRESPONDENCE ADDRESS:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                REGISTRATION NUMBER:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  APPLICATION NUMBER:
FILING DATE: 05-SEP
CLASSIFICATION: 536
                                                                                                                                                                     NUMBER OF SEQUENCES:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 OPERATING SYSTEM:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              amino acid
                                                                                                                                                                                                                                                                                 Washington
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         linear
                           GENERAL INFORMATION:
                                                                                                                                                                                                                                                                                                                                          USA
Patent No. 6280738
                                                                                                                                                                                                                                                                                                                                                                  20005
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    TELEPHONE:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         TOPOLOGY:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               US-08-923-992A-6
                                                                                                                                                                                                                                                                                                                                                                                                                                                            COMPUTER:
                                                                                                                                                                                                                                                                                                                                          COUNTRY:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      LENGTH:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     SEQUENCE
                                                                                                                                                                                                                                                                                                               STATE:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       q
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                òγ
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   q
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           ο̈́
```

379 QVTPKKRVKRDLAANENNQQKI 447 YHHNICEDPAHLREGTLDISE 1 1 1 422TAKSDSKTLDFSI 501 RDRRKHPRVWTLPPCPPDMDI : : : : 1 467 KNLKINESONYTL		ζο	110 QLEEDGIPVE
447 YHHHNICFDPAHLREGTLDISI 422TAKSDSKTTLDFSE 501 RDRRKHRPRVWTLPPCPPDMDI 467 KNLKLNESQYTL	EDVKFTV421	•	:
422TAKSDSKTTLDFSI 501 RDRRKHRPRVMTLPPCPPDMDI 1::::	PKLQERIANTWKRKGIKQKMHYSEPCDGAVTP 500	qa	169 EVEKD
501 RDRRKHRPRVMTLPPCPPDMDI :::::: 467 KNLKLNESOTYTL	: AEITI	Qy	169 -EPAATN- :
KNLKLNESQTVTL	MRTFKLPGFEKINDMVPYDRD 557	qq	210 EQKGITNEDK
		QY	222 PPIFSSRTCR
OY 558 DENRPAPPVKAPKKKGGKRKRTTDEEAAEPEEVDTAADDVKDAPEGPKEVPE	AADDVKDAPEGPKEVPE 610	QQ	
Db 514 SKTEEKVP-OEPKSNDKNOLQELIKSAQOELEKLEKAIKELMEQPEIPSN-PE	:: :	Oy Db	272 QELGLANARD : : 299 EELDKATTNE
RESULT 6 HS-031-0402-0		QY	
Sequence 2, Application US/08923992A Patent No. 6280738		qq	 351 KELKHNS
<pre>; GENERAL INFORMATION: ; APPLICANT: Tai, Joseph Y.</pre>		00y	392 OFCDKAATLE
Blake, Milan S. WENTION: No. 6280738-19A FC	Binding Forms of the Group B	q d	405 DFRKKFKTSE
Streptococcal Beta	3	Qy	437 EELNIPMVLD
<u> </u>	88 FOX P C . 1. 0	qq	458
STREET: 1100 New York Avenue, N.W., Suite 600	ite 600	ογ	491 SEPCDGAVTP
STATE: D.C. COUNTRY: USA		qq	498AEITI
ZIP: 20005 COMPUTER READABLE FORM:		, o	551 MVPYDRD
MEDIUM TYPE: Floppy disk COMPUTER: IBM PC compatible		qa	: 540 QVPKTPEQKD
OPERATING SYSTEM: PC_DOS/MS-DOS SOFTWARE: PatentIn Release #1.0, Version #1.30	# 1.30	Qy	608 VPE 610
CURRENT APPLICATION DATA: APPLICATION NUMBER: US/08/923,992A		qq	11 599 -PE 600
FILING DATE: 05-SEP-1997 CLASSIFICATION: 536 PRIOB ADDITCATION NAMES		RESULT	1LT 7
APPLICATION NUMBER: US 60/024,707		0-Sn	US-08-923-992A-4
FILING DATE: 06-SEP-1996 ATTORNEY/AGENT INFORMATION:		Pag.	Patent No. 6280738
NAME: Esmond, Robert W. REGISTRATION NUMBER: 32.893			GENERAL INFORMATION APPLICANT: Tai
REFERENCE/DOCKET NUMBER: 1438.0140001/RWE	SWE SWE		APPLICANT: Bla
TELECOMMUNICATION INFORMATION: TELEPHONE: (202) 371-2600		•	TITLE OF INV
TELEFAX: (202) 371-2540			CORRESPONDENCE
SEQUENCE CHARACTERISTICS:		••••	ADDRESSEE: S
min		•	Šr
MOLECULE TYPE: protein			COUNTRY: USA
US-U8-923-992A-2			COMPUTER READABL
Query Match 3.4%; Score 118.5; DB	DB 4; Length 1164;		COMPUTER: IBM OPERATING SYST
LOCAL SIMILATICY 18.4%; nes 122; Conservative 96)69; ; 256; Indels 189; Gaps 29;		SOFTWARE: Pa
QY 13 TPQSESSTFSSTLDSSAPSPARNLRRSGRNILQPSSEKDRDHEKRSGEELA- 	RDHEKRSGEELAGRM		APPLICATION NU FILING DATE: CLASSIFICATION
67MGKDANGHCLREGKEOEEGVKMAI	ERIAGEISAIDIGKKEKŲ LARMERRIORATKBOKK		PRIOR APPLICATION NO
			FILING DATE: ATTORNEY/AGENT

```
------KKAKQQKTLKQSDTKVDLSNIDKELNHQKSQVEKMA 209
                                                                                                                          ----IIEPEDAQDAAERGAARPPAVNSSYLPLPWKGRLGYACLNTYLRNAK 221
                                                                                                                                                       PSVVSRFPTAPYHHKSTNAEEREAKEPVLKTHSKDVEREAEIGVDDVVKM- 168
                                                                                                                                                                                                            RMASIVDHRHPLQFEDEPEHHLKNKPDKSKEPQ--DE-----LGHKFV 271
                                                                                                                                                                                                                                       GL------DQEIQEHVKKETSSEENTQKVDEHYANSLQNLAQKSL 298
                                                                                                                                                                                                                                                                                                                                    EQATOV---KNOFLENAOKLKEIOPLIKET--NVKLYKAMSESLEO---VE 350
NDVDNT1L.--SHEQKNEFKTKIDETNDSDALLELENQFNETNRLLHIKQHE 168
                                                                                                                                                                                                                                                                                              DIVKMLCWNEKYGIRFLRLSSEMFPFASHPVHGYKLAPFASEVLAEAGRVA 331
                                                                                                                                                                                                                                                                                                                                                                              HPGQFTQLGSPRKEVVESAIRDLEYHDELLSLLKLPEQQNRDAVMIIHMGG 391
                                                                                                                                                                                                                                                                                                                                                                                                               SEANLEDLVAKSKELVREYEGKLNQSKNLPELKQLEEERAHSKLKQVVE--- 404
                                                                                                                                                                                                                                                                                                                                                                                                                                                                  E-----RFKRNYA-----RLSQSCKNRLVLENDDVGWTVHDLLPVC 436
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             DYHHHNICFDPAHLREGTLDISD-----PKLQERIANTWKRKGIKQKMHY 490
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               PRDRRKHRPRVMTLPPCPPDMDLM1EAKDKEQAVFELMRTFKLPGFEKIND 550
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 DDENRPAPPVKAPKKKKGGKRKRTTDEEAAEPEEVDTAADDVKDAPEGPKE 607
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    ake, Milan S.
TTION: No. 6280738-1gA Fc Binding Forms of the Group B
TION: Streptococcal Beta Antigens
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  Sterne, Kessler, Goldstein & Fox P.L.L.C.
0 New York Avenue, N.W., Suite 600
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        BM PC compatible
STEM: PC-DOS/MS-DOS
Watchtin Release #1.0, Version #1.30
ATION DATA:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      NUMBER: US/08/923,992A
05-SEP-1997
ON: 536
ION DATA:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          APPLICATION NUMBER: US 60/024,707 FILING DATE: 06-SEP-1996 ATTORNEY/AGENT INFORMATION:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            cation US/08923992A
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                BLE FORM:
Floppy disk
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          Joseph Y.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  ADDRESS:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   ngton
```

```
; Sequence 11, Application PC/TUS9307261
; GENERAL INFORMATION:
; TITLE OF INVENTION:
; TITLE OF INVENTION:
                                                                                                                                                                                                                                                                                                                                                                                                                30;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        440 NIPMVLDYHHHNICFDPAHLREGTLDISD-----PKLQERIANTWKRKGIKQKMHYSEP 493
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    427 -----TAKSDSKTTLDFSDLITKYNPSVSDRISINYKINTDNHKI---- 466
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        494 CDGAVTPRDRRKHRPRVMTLPPCPPDMDLMIEAKDKEQAVFELMRTFKLPGFEKINDMVP 553
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               512 KTPEQKHSKTEQNVP----QEPKSNDKNQLQELIKSAQQELEKLEKAIKELMEQPEIPSN 567
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           320 KELKH---NSEANLQDLVAKSKEIVREYEGKLNQSKNLPELKQLEEEAHSKLKQVVEHFR 376
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       391 GQF--GDKAATLERFKRNYA------RLSQSCKNRLVLENDDVGWTVHDLLPVCEEL 439
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      377 KKFKTSEQVTPKKRVKRDLAANENNQQKIELTVSPENITVYEGEDVKFTV------ 426
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      467 --AEITIKNLKLNQSQTVTL-------KAKDDSGNVVE--KTFTITVQKKEEKQVP 511
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     554 -----YDRDDENRPAPPVKAPKKKGGKRKRTTDEEAAEPEEVDTAADDVKDAPEGPKE 607
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             138 EVEK------HNKPN-QQKTLKQSDTKVDLSNIDKELNHQKSQVEAMA 178
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           222 PPIFSSRTCRMASIVDHRHPLQFEDEPEHHLKNKPDKSKEPQ--DE-----LGHKFV 271
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        220 GKLFSSTKAGL------DQQIQEHVKKETSSEENTQKVDEHYANSLQNLAQKSL 267
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          272 QELGLANARDIVKMLCWNEKYGIRFLRLSSEMFPFASHPVHGYKLAPFASEVLAEAGRVA 331.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           268 EELDKATTNEQATQV---KNQFLENAQKLKEIQPLIKET--NVKLYKAMSESLEQ---VE 319
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          332 AELGHRLTTHPGQFTQLGSPRKEVVESAIRDLEYHDELLSLLKLPEQQNRDAVMII-HMG 390
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              67 -----MCKDANGHCLREGKEQEEGVKMAI-----EGLARMERRLQRATK----RQKK 109
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            110 QLEEDGIPVPSVVSRFPTAPYHHKSTNAEEREAKEPVLKTHSKDVEREAEIGVDDVVKME 169
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            170 PAA--TN-----IIEPEDAQDAAERGAARPPAVNSSYLPLPWKGRLGYACLNTYLRNAK 221
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             80 LQQWKNNLKNDVDNTIL--SHEQKNEFKTKIDETNDSDALLELENQFNETNRLHIKQHE 137
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           13 TPQSESSTFSSTLDSSAPSPARNLRRSGRNILQPSSEKDRDHEKRSGEELA---GRM--- 66
                                                                                                                                                                                                                                                                                                                                                         Query Match
3.4%; Score 117.5; DB 4; Length 1104;
Best Local Similarity 18.1%; Pred. No. 0.079;
Matches 120; Conservative 103; Mismatches 251; Indels 189;
                    REGISTRATION NUMBER: 32,893
REPERENCE/DOCKET NUMBER: 1438.0140001/RWE
TELECOMMUNICATION INFORMATION:
                                                                                     TELEPHONE: (202) 371-2600
TELEFAX: (202) 371-2540
INFORMATION FOR SEQ ID NO: 4:
SEQUENCE CHARACTERISTICS:
                                                                                                                                                                                                LENGTH: 1104 amino acids
Esmond, Robert W.
                                                                                                                                                                                                                                                                   MOLECULE TYPE: protein
                                                                                                                                                                                                                       amino acid
                                                                                                                                                                                                                                                    linear
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           608 VPE 610
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    PCT-US93-07261-11
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         568 -PE 569
                                                                                                                                                                                                                                            TOPOLOGY:
                                                                                                                                                                                                                                                                                             US-08-923-992A-4
                                                                                                                                                                                                                                                                                                                                                                      Query Match
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     RESULT
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            qq
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             δλ
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         q
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             οy
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              a
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               ŏ
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  qq
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  QQ
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                qq
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 δ
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 δλ
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           pp
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  ōλ
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               δy
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         Op
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                      δ
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    δ
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    δ
```

```
154 -----VEREAEIGVDDVVKMEPAATNIIEPEDAQDAAERGAARPPAVNSSYLPLPWKGR 207
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      223 IKDKGDGYEEIVETKFYGMRENALGELDEYEERYEKRYYLKEDGEGDLKDVEEKL---E 279
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   351 PRKEVVESA1RDL----EYHDELLSLLKLPEQQNRDAVMIIHMGGQFGDKAATLERFKRN 406
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 303 EQKKLKEDKEKKLIAAEEPDDE--KKIKLKDSDDKVVVPVNKNKSSFPDKFRAPDKKRTM 360
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  407 YARLSQSCKNRLVLENDDVGWTVHDLLPVCEELNIPMVLDYHHHNICFDPAHLREGTLDI 466
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 361 FYRLSELFP--IVPRKDNE-----LAVCGD-SMDSKVNGKKLKSTFNPFKRRNKL-- 408
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                41 RNILQPSSEKDRDHEKRSGEELAGRMMGKDANGHCLREGKEQEEGVKMAIEGLARMERRL 100
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               101 QRATKRQKKQLEEDGIPVPSVVSRFPTAPYHHKSTNAEEREAKEPV-----LKTHSKD 153
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   104 KDDKKIVPKKPESVEKDLKEMELKEKEFIK-QHLKDYEERKEKRRNWILRSLRRDKLREI 162
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              208 LGYACLNTYLRNAKPPIFSSRTCRMASIVDHRHPL------QFEDE-----PEHH 251
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            163 EQLEKLNAQLESAINELKERRASRRPMMVKMQRGMKDEVDEWIKKYDDEQAEKNGTKDEE 222
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          252 LKNKPDKSKE------PQDELG------HKFVQELGLANARDIVKMLCWNE 290
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         291 KYGIRFLRLSSEMFPFASHPVHGYKLAPFASEVLABAGRVAABLGHRLTTHPGQFTQLGS 350
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       280 ETGYGF----REKFP-----NTRILVKRKR-------NK 302
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 467 SDPKLQE--RIANTWKRKG1KQKMHYSEPCDGAVTPRDRRKHRPRVMTLPPCPPDMDLMI 524
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            18 RNALK---EKKLKEQKKNDAQKAKDLTKKES-----QDSSSEKSLKEKVNGEALKEKEN 68
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              Ouery Match
Best Local Similarity 19.4%; Pred. No. 0.14;
Matches 132; Conservative 108; Mismatches 247; Indels 193;
                                                                                                                                                                                                                                                                                                                                         APPLICATION NUMBER: PCT/US93/07261
FILING DATE: 19930805
                                                                                                                                                                                                                                                                                                                                                                FILING DATE: 19930805
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 07/927,531
FILING DATE: 07-AUG-1992
ATTORNEY/AGENT INFORMATION:
                                                                                                                                                                                                                                                                      OPERATING SYSTEM: Macintosh 6.0.5
                                                John H. C. Blasdale
                                                                                                                                                                                                                                                                                          Microsoft Word 5.1a
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           REGISTRATION NUMBER: 31,895
REFERENCE/DOCKET NUMBER: DX(TELECOMMUNICATION INFORMATION:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   INFORMATION FOR SEQ ID NO: 11:
                                                                                                                                                                                                                     MEDIUM TYPE: Floppy disk
COMPUTER: Apple Macintosh
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      NAME: Blascale, John H. C
REGISTRATION NUMBER: 31,89
                                                                     One Giralda Farms
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       1588 amino acids
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        201-822-7398
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           SEQUENCE CHARACTERISTICS:
                                                                                                                                                                                                                                                                                                               CURRENT APPLICATION DATA:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       201-822-7039
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            MOLECULE TYPE: protein
NUMBER OF SEQUENCES: 23
CORRESPONDENCE ADDRESS:
                                                                                                                                                                                              COMPUTER READABLE FORM:
                                                                                                                           New Jersey
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              amino acid
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          linear
                                                                                                                                                                    07940-1000
                                                                                                 Madison
                                                                                                                                                  HSA
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          TELEPHONE:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      PCT-US93-07261-11
                                                   ADDRESSEE:
                                                                                                                                                                                                                                                                                               SOFTWARE:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     TOPOLOGY:
                                                                                                                                                                                                                                               COMPUTER:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 TELEFAX:
                                                                                                                                             COUNTRY:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         LENGIH:
                                                                       STREET:
                                                                                                                        STATE:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           q
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    Op
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  QΥ
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   q
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   Qγ
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               QQ
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   δλ
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              90
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         ΟŊ
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     ^{\circ}
```

```
163 EQLEKLNAQLESAINELKERRASRRPMMVKMQRGMKDEVJEWIKKYDDEQAEKNGTKDEE 222
                                                                                                                                  -----HKFVQELGLANARDIVKMLCWNE 290
                                                                                                                                                             291 KYGIRFLRLSSEMFPFASHPVHGYKLAPFASEVLAEAGRVAAELGHRLTTHPGQFTQLGS 350
                                                                                                                                                                                                                                                                                                                   351 PRKEVVESAIRDL----EYHDELLSLLKLPEQQNRDAVMIIHMGGQFGDKAATLERFKRN 406
                                                                                                                                                                                                                                                                                                                                                             303 EOKKLKEDKEKKLIAAEEPDDE--KKIKLKDSDDKVVVPVNKNKSSFPDKFRAPDKKRIM 360
                                                                                                                                                                                                                                                                                                                                                                                                           407 YARLSQSCKNRLVLENDDVGWTVHDLLPVCEELNIPMVLDYHHHNICFDPAHLREGTLDI 466
                                                                                                                                                                                                                                                                                                                                                                                                                                              361 FYRESELFP--IVPRKDNE------LAVCGD-SMDSKVNGKKLKSTFNPFKRRNKL-- 408
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 ---QFEDE-----PEHH 251
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      467 SDPKLQE--RIANTWKRKGIKQKMHYSEPCDGAVTPRDRRKHRPRVMTLPPCPPDMDLMI 524
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     409 KERKMQELHKFKKNYKK---YQKLLEREKRE---NPDGEPLNTPEIHVIRP----SDLMD 458
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 525 EAKDKEQA-VFELMRTFKLPGFEKINDMVPYDRDDENRPAPPVKAPKKKKGG-KRKRTTD 582
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            583 EEAAEPEEVDTAADDVKDAPEGPKEVPEEERAMGGPYNRVYWPLGCEEWLKPKKREVKKG 642
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       APPLICANT: Weinzierl, Robert O.J.
TITLE OF INVENTION: TATA-BINDING PROTEIN ASSOCIATED FACTORS,
TITLE OF INVENTION: NUCLEIC ACIDS ENCODING TAFS AND METHODS OF USE
                                                                                                                                                                                                                                                                  280 ETGYGF----REKEP-----TTRILVKRKR------
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             ADDRESSEE: FLEHR, HOHBACH, TEST, ALBRITTON & HERBERT STREET: 4 Embarcadero Center, Suite 3400
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      SOFTWARE: Patentin Release #1.0, Version #1.25 CURRENT APPLICATION DATA:
                                  208 LGYACLNTYLRNAKPPIFSSRTCRMASIVDHRHPL--
                                                                                                                             252 LKNKPDKSKE------PQDELG-----
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  APPLICATION NUMBER: US/08/188,582
FILING DATE: 28-JAN-1994
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                PC-DOS/MS-DOS
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       ; Sequence 14, Application US/08188582
; Patent No. 5534410
; GENERAL INFORMATION:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     APPLICANT: Dynlact, Brian D. APPLICANT: Hoey, Timothy APPLICANT: Ruppert, Siegfried APPLICANT: Tanese, Naoko
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       IBM PC compatible
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             NAME: Osman, Richard A
REGISTRATION NUMBER: 36,627
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          643 KVPEE--VEDEG----EFD 655
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      550 HVSREYQLDNEGPSTLKEYD 569
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       ATTORNEY/AGENT INFORMATION:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         APPLICANT: Tjian, Robert
APPLICANT: Comai, Lucio
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     Wang, Edith
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         San Francisco
California
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         CORRESPONDENCE ADDRESS
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   NUMBER OF SEQUENCES:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             OPERATING SYSTEM:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                94111-4187
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               CLASSIFICATION:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            USA
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       COMPUTER:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 US-08-188-582-14
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        APPLICANT:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          COUNTRY:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    STATE:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        RESULT 10
                                  ò
                                                                            g
                                                                                                                             ò
                                                                                                                                                                    QQ
                                                                                                                                                                                                                     Ωÿ
                                                                                                                                                                                                                                                               qq
                                                                                                                                                                                                                                                                                                                ά
                                                                                                                                                                                                                                                                                                                                                             q
                                                                                                                                                                                                                                                                                                                                                                                                                                                     QQ
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      δ
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 Óγ
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             QΩ
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       ŏ
                                                                                                                                                                                                                                                                                                                                                                                                                                           TITLE OF INVENTION: PFEMP3 MALARIA ANTIGEN, ANALOGS, ANTIBODIES AND USES THEREOF
409 KERKMQELHKEKKNYKK---YQKLLEREKRE---NPDGEPLNTPEIHVIRP----SDLMD 458
                                            525 EAKDKEQA-VFELMRTFKLPGFEKINDMVPYDRDDENRPAPPVKAPKKKGG-KRKRTTD 582
                                                                                      459 KGENKSAGHPFKYQPTKGLKEYEESHVSKDYQLEHE----PPTKLPEYEKGHVSREYQLD 514
                                                                                                                                       583 EEAAEPEEVDTAADDVKDAPEGPKEVPEEERAMGGPYNRVYWPLGCEEWLKPKKREVKKG 642
                                                                                                                                                                            ------PTKLPEYEK---GHVSREY---QLDNEVRDELPEYEKG 549
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          154 -----VEREAEIGVDDVVKMEPAATNIIEPEDAQDAAERGAARPPAVNSSYLPLPWKGR 207
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        41 RNILQPSSEKDRDHEKRSGEELAGRMMGKDANGHCLREGKEQEEGVKMAIEGLARMERRL 100
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            18 RNALK---EKKLKEOKKNDAOKAKDLTKKES-----ODSSSEKSLKEKVNGEALKEKEN 68
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     Ouery Match 3.4%; Score 117.5; DB 5; Length 1663; Best Local Similarity 19.4%; Pred. No. 0.15; Matches 132; Conservative 108; Mismatches 247; Indels 193;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             101 QRATKRQKKQLEEDGIPVPSVVSRFPTAPYHHKSTNAEEREAKEPV---
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      COMPUTER: Apple Macintosh
OPERATING SYSTEM: Macintosh 6.0.5
SOFTWARE: Microsoft Word 5.1a
CURRENT APPLICATION DATA:
APPLICATION NUMBER: PCT/US93/07261
FILING DATE: 19930805
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             PRIOR APPLICATION DAYS:
APPLICATION NUMBER: US 07/927,531
FILING DATE: 07-AUG-1992
ATTORNEY/AGENT INFORMATION:
                                                                                                                                                                                                                                                                                                                                                                                                 Sequence 16, Application PC/TUS9307261 GENERAL INFORMATION:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    Plasmodium falciparum
                                                                                                                                                                                                                                                                                                                                                                                                                                                                 NUMBER OF SEQUENCES: 23
CORRESPONDENCE ADDRESS:
ADDRESSEE: John H. C. Blasdale
STREET: One Giralda Farms
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           NAME: Blasdale, John H. C.
REGISTRATION NUMBER: 31,895
REFRENCE/DOCKET NUMBER: DX028
TELECOMMUNICATION INFORMATION:
TELEPHONE: 201-822-7398
TELEFAX: 201-822-7399
                                                                                                                                                                                                                                  643 KVPEE--VEDEG----EFD 655
                                                                                                                                                                                                                                                                              550 HVSREYQLDNEGPSTLKEYD 569
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         TELEFAX: 201-822-7039
INFORMATION FOR SEQ ID NO: 16:
SEQUENCE CHARACTERISTICS:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                MEDIUM TYPE: Floppy disk
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   LENGTH: 1663 amino acids
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     MOLECULE TYPE: peptide ORIGINAL SOURCE:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            COMPUTER READABLE FORM:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          ; STRAIN: Malayan Camp
PCT-US93-07261-16
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            CITY: Madison
STATE: New Jersey
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            amino acid
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         COUNTRY: USA
ZIP: 07940-1000
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       linear
                                                                                                                                                                                     515 HEP-----
                                                                                                                                                                                                                                                                                                                                                                         PCT-US93-07261-16
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             TOPOLOGY:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      ORGANISM:
                                                                                                                                                                                                                                                                              qq
                                              δ
                                                                                        a
                                                                                                                                         ò
                                                                                                                                                                                   셤
                                                                                                                                                                                                                               ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                ô
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        ò
```

Db 1757 ----IRPKQPRMLQENTRMDMENE 1776

```
1596 IVNVCYQTLTEYDEH------LTQLEKDICTAKEAA-LEEAELESLDPMTPGPYTPQP 1646
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       1078 SSTEVLSTDTDSSSAEDS-DFEEMGKNIENMLQNKKTSSQLSREREEQERKELQRMLLAA 1136
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      1137 GSAASGNNHRDDDTASVTSLNSSATGRCLKIYRTFRDEEGKEYVRCETVRKPAVIDAYVR 1196
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        1197 IRTTKDEEFIRKFALFDEQHREEMRKERRRIQEQLRRLKRNQEKEKLKGPPEKKPKKMKE 1256
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     : | | : : | | : : | | 1257 RPDLKLKCGACGAIGHMRTNKFCPLYYQ-TNA---PPSNPVAMTEEQEEELEKTVIHNDN 1312
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 1313 EELIKVE--GTKIVLGKOLIESADEVRRKSLVLKFPKQQLPPKKKRRVGTTVHCDYLNRP 1370
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         1371 HKSIHRRRIDPMVILSSILESIINDMRDLPNIYPFHTPVNAKVVKDY--YKIITRPMDLQ 1428
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        1429 TLRENVRK------RLYPSREEFREH------LELIVK-----N 1455
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        1456 SATYNG------PKHSLTQISQSMLDLCDE-----KLKEKE------DKL 1488
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       1538 PDY--YKVIVNPMDLETIRKNISKHKYQSRESFLDDVNLILANSVKYNGPESQYTKTAQE 1595
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              1647 PDLYDINTSLSMSRDASVFQDESNMSVLDIPSATPEKQVTQEGEDGDGDLADEEEGTVQQ 1706
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                1707 PQASVLYEDLLMSEGEDDEEDAGSDEEGDNPFSAIQLSESGSDSDVGSGG-------- 1756
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  398 ATLERFKRNYARLSQSCKNRLVLENDDVGWT------VHDLLPVCEELNIPMV 444
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        445 LDYHHHNICFDPAHLREGTLDISDPKLQER-----IANTWKRKG----IKQKMH 489
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    490 YSEPCDGAVTPRDRRKHRPRVMTLPPCPPDMDLMIEAKDKEQAVFELM-----RTFKL 542
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                120 -----SVVSRFPT---APYHHKSTNAEEREAKEPVLKTHSKDVEREAEI---GV 162
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               163 DDVVKMEPAATNIIEPEDAQDAAERGAARPPAVN--SSYLPLPWKGRLGYACLNTYLRNA 220
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  221 KPPIFSSRTCRMASIVDHRHPL--QFEDEPEHHLKNKPDKSKEPQDELGHKFV-QELGLA 277
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 278 NARDIVKMLCWNEKYGIRFLRLSSEMFPFASHPVHGYKLAPFASEVLAEAGRVAAELGHR 337
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   338 LTTHPGQFTQLGSPRKEVVESAIRDLEYHDELLSLLKLPEQQNRDAVMIIHMGGQFGDKA 397
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     ----AEPEEVDTAADDVKDAPEGPKEVPE----EERAMGGPYNRVYWPLG 627
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              82 -----QEEGV-KMAI------EGLARMERRLQRATKRQKKQLEEDGIPVP----- 119
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              Matches 149; Conservative 107; Mismatches 275; Indels 273; Gaps
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            16 SESSTFSSTLDSSAPSPARNLRRSGRNI-----LQPSSEKDRDHEKRSGEEL-----A 63
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      543 PGFEKINDMVPYDRD-----DENRPAP---PVKAPKKK-----KGGKRKRTTDEEA---
                                                                                                                                                                                                                                                                                                                                                                                                                                       Ouery Match 3.4%, Score 117.5; DB 1; Length 1872; Best Local Similarity 18.5%; Pred. No. 0.18;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              64 G------EGKE----
REFERENCE/DOCKET NUMBER: A-57650-2/AJT/RAO
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        628 CEEWLKPKKREVKKGKVPEEVEDE 651
                                                       TELEPHONE: (415) 781-1989
TELEFAX: (415) 388-3249
TELERAX: 910 277299
INFORMATION FOR SEQ ID NO: 14:
SEQUENCE CHARACTERISTICS:
LENGTH: 1872 amino acids
                                 TELECOMMUNICATION INFORMATION:
                                                                                                                                                                                                                                                                      single
                                                                                                                                                                                                                                                                                                                              MOLECULE TYPE: peptide
                                                                                                                                                                                                                                          TYPE: amino acid
STRANDEDNESS: sir
                                                                                                                                                                                                                                                                                                  linear
                                                                                                                                                                                                                                                                                               TOPOLOGY:
                                                                                                                                                                                                                                                                                                                                                          US-08-188-582-14
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                qq
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               q
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           q
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    δ
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   οy
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           рp
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  δλ
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         qq
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    δy
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           a
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      δλ
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        Dp
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     δ
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        òχ
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          δ
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            ŏ
```

```
38:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         1078 SSTEVLSTDTDSSSAEDS-DFFEEMGKNIENMLQNKKTSSQLSREREEQERKELQRMLLAA 1136
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               1137 GSAASGNNHRDDDTASVTSLNSSATGRCLKIYRTFRDEEGKEYVRCETVRKPAVIDAYVR 1196
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   1197 IRTTKDEEFIRKFALFDEQHREEMRKERRRIQEQLRRLKRNOEKEKLKGPPEKKPKKMKE 1256
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        82 -----QEEGUV-KMAI-----EGLARMERRLQRATKRQKKQLEEDGIPVP------ 119
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 120 -----SVVSRFPT---APYHHKSTNAEEREAKEPVLKTHSKDVEREAEI---GV 162
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                Query Match 3.4%; Score 117.5; DB 1; Length 1872; Best Local Similarity 18.5%; Pred. No. 0.18; Matches 149; Conservative 107; Mismatches 275; Indels 273; Gaps
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           16 SESSTFSSTLDSSAPSPARNLRRSGRNI-----LQPSSEKDRDHEKRSGEEL-----A 63
                                                                                                                                                                                                                              APPLICANT: Wang, Edith
APPLICANT: Weinzierl, Robert O.J.
TITLE OF INVENTION: TATA-BINDING PROTEIN ASSOCIATED FACTORS,
TITLE OF INVENTION: NUCLEIC ACIDS ENCODING TAFS AND METHODS OF USE
NUMBER OF SEQUENCES: 36
                                                                                                                                                                                                                                                                                                                                                                                 4 Embarcadero Center, Suite 3400
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    64 G------EGKE----
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           PatentIn Release #1.0, Version #1.25
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             A-57650-2/AJT/RAO
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               UMBER: US/08/646,715
09-MAY-1996
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          APPLICATION NUMBER: US 08/188,582
FILING DATE: 28-JAN-1994
ATTORNEY/AGENT INFORMATION:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            PC-DOS/MS-DOS
                                      Sequence 14, Application US/08646715 Patent No. 5637686
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          REFERENCE/DOCKET NUMBER: A-E
TELECOMMUNICATION INFORMATION:
TELEPHONE: (415) 781-1989
TELEFAX: (415) 398-3249
                                                                                                                                                    Dynlact, Brian D.
Hoey, Timothy
Ruppert, Siegfried
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  IBM PC compatible
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       36,627
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   TELEX: 910 277299
INFORMATION FOR SEQ ID NO: 14:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            Floppy disk
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        1872 amino acids
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  NAME: Osman, Richard A REGISTRATION NUMBER: 36
                                                                               GENERAL INFORMATION:
APPLICANT: Tjian, Robert
APPLICANT: Comai, Lucio
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                CURRENT APPLICATION DATA:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   SEQUENCE CHARACTERISTICS:
                                                                                                                                                                                                                            Tanese, Naoko
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 CLASSIFICATION: 435
PRIOR APPLICATION DATA:
APPLICATION NUMBER: U
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      single
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            ; MOLECULE TYPE: peptide US-08-646-715-14
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       COMPUTER READABLE FORM:
                                                                                                                                                                                                                                                                                                                                                                CORRESPONDENCE ADDRESS:
                                                                                                                                                                                                                                                                                                                                                                                                                              San Francisco
California
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             APPLICATION NUMBER:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                amino acid
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       OPERATING SYSTEM:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              inear
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               94111-4187
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      STRANDEDNESS:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               MEDIUM TYPE:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   FILING DATE:
                                                                                                                                                                                                                                                                                                                                                                                     ADDRESSEE:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       COMPUTER:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              TOPOLOGY:
                   US-08-646-715-14
                                                                                                                                                                                                  APPLICANT:
                                                                                                                                                         APPLICANT:
                                                                                                                                                                              APPLICANT:
                                                                                                                                                                                                                            APPLICANT:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                              COUNTRY:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      LENGTH:
                                                                                                                                                                                                                                                                                                                                                                                                             STREET:
                                                                                                                                                                                                                                                                                                                                                                                                                                 CITY:
STATE:
RESULT 11
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  q
```

```
1257 RPDLKLKCGACGAIGHMRTNKFCPLYYQ-TNA---PPSNPVAMTEGQEEELEKTVIHNDN 1312
                                                                                                                   1313 EELIKVE--GTKIVLGKQLIESADEVRRKSLVLKFPKQQLPPKKKRRVGTTVHCDYLNRP 1370
                                                                                                                                                                                                                   1371 HKSIHRRRIDPMVTLSSILESIINDMRDLPNTYPFHTPVNAKVVKDY --YKIITRPMDLQ 1428
                                                                                                                                                                                                                                                                                                                    1429 TLRENVRK-------RLYPSREEFREH------LELIVK-----N 1455
                                                                                                                                                                                                                                                                                                                                                                                                                     1456 SATYNG------PKHSLTQISQSMLDLCDE-----KLKEKE-------DKL 1488
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         1647 PDLYDTNTSLSMSRDASVFQDESNMSVLDIPSATPEKQVTQEGEDGDGDLADEEEGTVQQ 1706
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       1596 IVNVCYQTLTEYDEH-----LTQLEKDICTAKEAA-LEEAELESLDPMTPGPYTPQP 1646
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         1707 PQASVLYEDLLMSEGEDDEEDAGSDEEGDNPFSAIQLSESGSDSDVGSGG------ 1756
                                                                      163 DDVVKMEPAATNIIEPEDAQDAAERGAARPPAVN--SSYLPLPWKGRLGYACLNTYLRNA 220
                                                                                                                                                                       221 KPPIFSSRTCRMASIVDHRHPL--QFEDEPEHHLKNKPDKSKEPQDELGHKFV-QELGLA 277
                                                                                                                                                                                                                                                                     278 NARDIVKMLCWNEKYGIRFLRLSSEMFPFASHPVHGYKLAPFASEVLAEAGRVAAELGHR 337
                                                                                                                                                                                                                                                                                                                                                                   338 LTTHPGQFTQLGSPRKEVVESAIRDLEYHDELLSLLKLPEQQNRDAVMIIHMGGQFGDKA 397
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       445 LDYHHHNICFDPAHLREGTLDISDPKLQER-----IANTWKRKG-----IKQKMH 489
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     490 YSEPCDGAVTPRDRRKHRPRVMTLPPCPPDMDLMIEAKDKEQAVFELM-----RTFKL 542
                                                                                                                                                                                                                                                                                                                                                                                                                                                                         398 ATLERFKRNYARLSQSCKNRLVLENDDVGWT------VHDLLPVCEELNIPMV 444
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           APPLICANT: Wang, Edith
APPLICANT: Weinzierl, Robert O.J.
TITLE OF INVENTION: TATA-BINDING PROTEIN ASSOCIATED FACTORS,
TITLE OF INVENTION: NUCLEIC ACIDS ENCODING TAFS AND METHODS OF USE
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         543 PGFEKINDMVPYDRD----DENRPAP---PVKAPKK-----KGGKRKRTTDEEA----
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            --------AEPEEVDTAADDVKDAPEGPKEVPE----EERAMGGPYNRVYWPLG
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                ADDRESSEE: FLEHR, HOHBACH, TEST, ALBRITTON & HERBERT STREET: 4 Embarcadero Center, Suite 3400 CITY: San Francisco STATE: California COUNTRY: USA
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      SOFTWARE: PatentIn Release #1.0, Version #1.25 CURRENT APPLICATION DATA:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            1757 ---- IRPKOPRMLOENTRMDMENE 1776
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    Sequence 11, Application US/08188582
Patent No. 5534410
GENERAL INFORMATION:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          628 CEEWLKPKKREVKKGKVPEEVEDE 651
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              APPLICANT: Ruppert, Siegfried APPLICANT: Tanese, Naoko
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        APPLICANT: Tjian, Robert
APPLICANT: Comai, Lucio
APPLICANT: Dynlact, Brian D.
APPLICANT: Hoey, Timothy
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   NUMBER OF SEQUENCES: 36
CORRESPONDENCE ADDRESS:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy of
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             94111-4187
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               US-08-188-582-11
                                                                         ó
                                                                                                                                                                                                                   g
                                                                                                                                                                                                                                                                                                                                                                                                                     g
                                                                                                                                                                                                                                                                        ò
                                                                                                                                                                                                                                                                                                                    qq
                                                                                                                                                                                                                                                                                                                                                                   ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                         õ
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                q
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       Q
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         QQ
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         ag
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             ò
```

```
38;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           1099 SSTEVLSTDTDSSSAEDS-DFEEMGKNIENMLQNKKTSSQLSREREEQERKELQRMLLAA 1157
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     : | | ::: | | | :178 RPDLKLKCGACGAIGHMFTNKFCPLYYQ-TNA---PPSNPVAMTEEQEEELEKTVIHNDN 1333
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          1158 GSAASGNNHRDDDTASVTSLNSSATGRCLKIYRTFRDEEGKEYVRCETVRKPAVIDAYVR 1217
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            1334 BELIKVE--GTKIVLGKQLIESADEVRRKSLVLKFPKQQLPPKKKRRVGTTVHCDYLNRP 1391
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         1392 HKSIHRRTDPMVTLSSILESIINDMRDLPNTYPFHTPVNAKVVKDY--YKIITRPMDLQ 1449
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        1450 TLRENVRK-------RLYPSREEFREH-------LELIVK------N 1476
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   1477 SATYNG------PKHSLTQISQSMLDLCDE-----KLKEKE-------DKL 1509
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        1559 PDY -- YKVIVNPMDLETIRKNISKHKYQSRESFLDDVNLILANSVKYNGPESQYTKTAQE 1616
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 1617 IVNVCYQTLTEYDEH-----LTQLEKDICTAKEAA-LEEAELESLDPMTPGPYTPQP 1667
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          82 -----QEEGV-KMAI-----EGLARMERRLQRATKRQKKQLEEDGIPVP----- 119
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    120 -----SVVSRFPT---APYHHKSTNAEEREAKEPVLKTHSKDVEREAEI---GV 162
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 163 DDVVKMEPAATNIIEPEDAQDAAERGAARPPAVN--SSYLPLPWKGRLGYACLNTYLRNA 220
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            221 KPPIFSSRTCRMASIVOHRHPL--QFEDEPEHHLKNKPDKSKEPQDELGHKFV-QELGLA 277
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        338 LTTHPGQFTQLGSPRKEVVESAIRDLEYHDELLSLLKLPEQQNRDAVMIIHMGGQFGDKA 397
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     398 ATLEREKRNYARLSQSCKNRLVLENDDVGWT------VHDLLPVCEELNIPMV 444
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            1668 PDLYDTNTSLSMSRDASVFQDESNMSVLDIPSATPEKOVTQEGEDGDGDLADEEEGTVQQ 1727
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         278 NARDIVKMLCWNEKYGIRFLRLSSEMFPFASHPVHGYKLAPFASEVLAEAGRVAAELGHR 337
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     445 LDYHHHNICFDPAHLREGTLDISDPKLQER-----IANTWKRKG----IKQKMH 489
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    490 YSEPCDGAVTPRDRRKHRPRVMTLPPCPPDMDLMIEAKDKEQAVFELM-----RTFKL 542
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 543 PGFEKINDMVPYDRD-----DENRPAP---PVKAPKKK-----KGGKRKRTTDEEA---- 585
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           16 SESSTFSSTLDSSAPSPARNLRRSGRNI-----LQPSSEKDRDHEKRSGEEL-----A 63
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              Indels 273;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     DB 1; Length 1893;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             64 G-----EGKE-----RMMGKDANGHCLR------EGKE-----
                                                                                                                                                                                                                                                                                                                                                                                                                                                                              Query Match 3.4%; Score 117.5; DB 1; Best Local Similarity 18.5%; Pred. No. 0.18; Matches 149; Conservative 107; Mismatches 275;
                                                                                                                                               A-57650-2/AJT/RAO
APPLICATION NUMBER: US/08/188,582
FILING DATE: 28-JAN-1994
CLASSIFICATION: 435
                              Filther
CLASSIFICATION: 4.5.
ATTORNEY/AGENT INFORMATION:
NAME: OSMAN, Richard A
                                                                                                                                          REFERENCE/DOCKET NUMBER: A-TELECOMMUNICATION:
                                                                                                                                                                                       TELEPHONE: (415) 781-1989
TELEFAX: (415) 398-3249
TELEX: 910 277299
INFORMATION FOR SEQ ID NO: 11:
SEQUENCE CHARACTERISTICS:
                                                                                                                                                                                                                                                                                                                1893 amino acids
                                                                                                                                                                                                                                                                                                                                                                                   ; MOLECULE TYPE: protein US-08-188-582-11
                                                                                                                                                                                                                                                                                                                                      amino acid
                                                                                                                                                                                                                                                                                                                                                               linear
                         FILING DATE: 28 CLASSIFICATION:
                                                                                                                                                                                                                                                                                                                                                             TOPOLOGY:
                                                                                                                                                                                                                                                                                                                LENGTH:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              586
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             δλ
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        Q
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    qq
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    Óγ
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               Dp
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              qq
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              ōλ
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         Вb
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     qq
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         οy
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        ó
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  QC
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               qq
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               Qγ
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           δλ
```

```
MEDIUM TYPE:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           94105
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        ADDRESSEE:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   CA
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  SOFTWARE:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   US-08-475-844-9
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         COUNTRY:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   STATE:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 RESULT.
                                                                                                                                                                                                                           ŏ
                                                                                                                                                                                                                                                                                                                                                                                           ^{\circ}
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          38;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          1158 GSAASGNNHRDDDTASVTSLNSSATGRCLKIYRTFRDEEGKEYVRCETVRKPAVIDAYVR 1217
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       1099 SSTEVLSTDTDSSSAEDS-DFEEMCKNIENMLQNKKTSSQLSREREEQERKELQRMLLAA 1157
1728 PQASVLYEDLLMSFGEDDEEDAGSDEEGDNPFSAIQLSESGSDSDVGSGG------ 1777
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  82 -----QEEGV-KMAI-----EGLARMERRLQRATKRQKKQLEEDGIPVP------ 119
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 16 SESSTFSSTLDSSAPSPARNLRRSGRNI-----LQPSSEKDRDHEKRSGEEL-----A 63
                                                                                                                                                                                                                                                                                                                                                                                                APPLICANT: Wang, Edith
APPLICANT: Weinzierl, Robert O.J.
ITILE OF INVENTION: TATA-BINDING PROTEIN ASSOCIATED FACTORS,
TITLE OF INVENTION: NUCLEIC ACIDS ENCODING TAFS AND METHODS OF USE
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          Matches 149; Conservative 107; Mismatches 275; Indels 273;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              3.4%; Score 117.5; DB 1; Length 1893; 18.5%; Pred. No. 0.18;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     -----RMMGKDANGHCLR------EGKE------
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      E: FLEHR, HOHBACH, TEST, ALBRITTON & HERBERT 4 Embarcadero Center, Suite 3400
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        OPERATING SYSTEM: PC_DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               REFERENCE/DOCKET NUMBER: A-57650-2/AJT/RAO TELECOMMUNICATION INFORMATION:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           CLASSIFICATION: 435
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/188,582
FILING DATE: 28-JAN-1994
ATTORNEY/AGENT INFORMATION:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     APPLICATION NUMBER: US/08/646,715
FILING DATE: 09-MAY-1996
                                                                                                       1778 ----IRPKOPRMLOENTRMDMENE 1797
                                                                 628 CEEWLKPKKREVKKGKVPEEVEDE 651
                                                                                                                                                                                                         Sequence 11, Application US/08646715 Patent No. 5637686 GENERAL INFORMATION:
                                                                                                                                                                                                                                                                                                                 Dynlact, Brian D.
Hoey, Timothy
Ruppert, Siegfried
Tanese, Naoko
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              'AGENT INFOLL
OSman, Richard A
... with RER: 36,627
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         IBM PC compatible
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                TELERS: (415) 398-3249
TELEX: 910 277299
INFORMATION FOR SEQ ID NO: 11:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   (415) 781-1989
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        LENGTH: 1893 amino acids
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    SEQUENCE CHARACTERISTICS:
                                                                                                                                                                                                                                                                         Tjian, Robert
                                                                                                                                                                                                                                                                                                 Lucio
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   ) MOLECULE TYPE: protein US-08-646-715-11
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     CORRESPONDENCE ADDRESS: ADDRESSE: FLEHR, HO
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      REGISTRATION NUMBER:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    San Francisco
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   NUMBER OF SEQUENCES:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      California
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     Query Match
Best Local Similarity
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              amino acid
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    linear
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             94111-4187
                                                                                                                                                                                                                                                                                                 Comai,
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     TELEPHONE:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                ropology:
                                                                                                                                                                                          US-08-646-715-11
                                                                                                                                                                                                                                                                           APPLICANT:
                                                                                                                                                                                                                                                                                                                                    APPLICANT:
APPLICANT:
                                                                                                                                                                                                                                                                                                                                                                                                APPLICANT:
APPLICANT:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            COUNTRY:
                                                                                                                                                                                                                                                                                                 APPLICANT:
                                                                                                                                                                                                                                                                                                                                                                              APPLICANT:
                                                                                                                                                                                                                                                                                                                   APPLICANT:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    CITY: STATE:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 STREET:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       64 G---
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       ó
                            qq
                                                                    ŏ
```

```
1477 SATYNG------PKHSLTQISQSMLDLCDE-----KLKEKE-------DKL 1509
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    1510 ARLEK------AINPLLDDDDQVAFSFILDNIVTQKMMAVPDSWPFHHPVNKKFV 1558
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                1559 PDY--YKVLVNPMDLETIRKNISKHKYOSRESFLDDVNLILANSVKYNGPESQYTKTAQE 1616
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       1617 IVNVCYQTLTEYDEH------LTQLEKDICTAKEAA-LEEAELESLDPMTPGPYTPQP 1667
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   1668 PDLYDTNTSLSMSRDASVFQDESNMSVLDIPSATPEKQVTQEGEDGDGDLADEEEGTVQQ 1727
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           1728 PQASVLYEDLIMSEGEDDEEDAGSDEEGDNPFSAIQLSESGSDSDVGSGG------ 1777
                                                                                                                                               1278 RPDLKLKCGACGAIGHMRTNFFCPLYYQ-TNA---PPSNPVAMTEEQEEELEKTVIHNDN 1333
                                                                                                                                                                                                                                                                                                                                                                                                                                                         1392 HKSIHRRRTDPMVTLSSILESILNDMRDLPNTYPFHTPVNAKVVKDY--YKLITRPMDLQ 1449
                          1218 IRTTKDEEFIRKFALFDEQHREEMRKERRIQEQLRRLKRNQEKEKLKGPPEKKPKKMKE 1277
                                                                                                                                                                                                                                                                                                                     1334 EELIKVE--GTKIVLGKOLIESADEVRRKSLVLKFPKOQLPPKKKRRVGTTVHCDYLNRP 1391
                                                                                                                                                                                                                                                                                                                                                                                 221 KPPIFSSRTCRMASIVDHRHPL--QFEDEPEHHLKNKPDKSKEPQDELGHKFV-QELGLA 277
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 278 NARDIVKMLCWNEKYGIRFLRLSSEMFPFASHPVHGYKLAPFASEVLAEAGRVAAELGHR 337
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    398 ATLERFKRNYARLSQSCKNRLVLENDDVGWT------VHDLLPVCEELNIPMV 444
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                445 LDYHHHNICFDPAHLREGTLDISDPKLQER------IANTWKRKG----IKQKMH 489
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       490 YSEPCDGAVTPRDRRKHRPRVMTLPPCPPDMDLMIEAKDKEQAVFELM-----RTFKL 542
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               543 PGFEKINDMVPYDRD----DENRPAP---PVKAPKKK-----KGGKRKRTTDEEA---- 585
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   586 ------EERAMGGPYNRVYWPLG 627
                                                                                                                                                                                                                                             163 DDVVKMEPAATNIIEPEDAQDAAERGAARPPAVN--SSYLPLPWKGRLGYACLNTYLRNA 220
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        338 LTTHPGQFTQLGSPRKEVVESAIRDLEYHDELLSLLKLPEQQNRDAVMIIHMGGQFGDKA 397
                                                                                                     120 ----SVVSRFPT---APYHHKSTNAEEREAKEPVLKTHSKDVEREAEI---GV 162
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       Townsend and Townsend Khourie and Crew
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              Patentin Release #1.0, Version #1.25
COMPUTER: IBM PC compatible OPERATING SYSTEM: PC-DOS/MS-DOS
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           1778 ----IRPKQPRMLQENTRMDMENE 1797
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      628 CEEWLKPKKREVKKGKVPEEVEDE 651
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          APPLICANT: Lobanenkov, Victor V. APPLICANT: Neiman, Paul B. APPLICANT: Klenova, Elena M. APPLICANT: Goodwin, Graham H.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        Sequence 9, Application US/08475844 Patent No. 5972643 GENERAL INFORMATION:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  APPLICANT: Filippova, Galina N. APPLICANT: Collins, Steven J. TITLE OF INVENTION: CTCF
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            One Market Plaza
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                Floppy disk
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      COMPUTER READABLE FORM:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   NUMBER OF SEQUENCES: 2
CORRESPONDENCE ADDRESS:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              San Francisco
   -:-
```

```
238 SEVNAEKVVGNMKPPRPTKIKKKGVKKTFQCELCSYTCPRRSNLD-RHMKSHTDERPHKC 296
                                                                        EAAEPEEVDTAADDVKDAPEGP-----KEVPEEERAMGGPYNRVYWPLGCEEWLKPK 635
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              135 TNAEEREAKEPVI,---KTHSKD--VEREAEIGVDD------VVKME------------- 169
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       58 VOMVMMEQLDPTLLOMKTEVMEGTVAPEAEAAVDDTQIITLQVVNMEEQPINIGELQLVQ 117
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            118 VPVPVTVPVATTSVEELQGAYENEVSKEGLAESEPMICHTLPLPEGFQVVKVGANGEVET 177
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              178 LEGGELPPQEDPSWQKDPDYQPPAKKTKKTKKSKLRYTEEGKDVDVSVYDFEEEQQEGLL 237
                                                                                                                         663 KONOP----TALIQVEDONTGAIENIIVEVKKEPDAEPAEGE------EEEAQPA 707
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         84 EGVKMAIEGLARMERRLQRATKRQKKQ--LEEDGIPVP-----SVVSRFPTAPYHHKS 134
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              251 -------HLKNKPDKSKEPQDELGHKFVQELGLANARDIVKMLCWNEK 291
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       4 DAVEAIVEESETFIKGKERKTYQRRREGGGEBDACHLPQNQTDGGEVVQDV----NSS 57
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               --IFSSRTCRMASIVDHRHPLQFEDEPEH--
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             204 -----ACLNTY------
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          Indels 304;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       Query Match 3.3%; Score 115; DB 5; Length 727; Best Local Similarity 17.7%; Pred. No. 0.07; Matches 140; Conservative 97; Mismatches 252; Indels 3
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             170 ----PAATNIIEPEDA---QDAAERGAARPPAVNSSYLPLP----
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     SOFTWARE: Patentin Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: PCT/US95/08429
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 14538A-11-1PC
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/261,680
FILING DATE: 17-JUN-1994
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC COMPATIBLE
OPERATING SYSTEM: PC-DOS/MS-DOS
                                                                                                                                                                                                                                                                                                                                       Sequence 9, Application PC/TUS9508429 GENERAL INFORMATION:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               REGISTRATION NUMBER: 31,990
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              REFERENCE/DOCKET NUMBER: 14 TELECOMMUNICATION:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    TELEFAX: 415-543-5043
1NFORMATION FOR SEQ ID NO: 9:
SEQUENCE CHARACTERISTICS:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  Steven W.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   ATTORNEY/AGENT INFORMATION:
NAME: Parmelee, Steven W
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     15-JUN-1995
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               206-467-9600
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  727 amino acids
                                                                                                                                                                                                                                                                                                                                                                                                                          CTCF
21
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       ; MOLECULE TYPE: protein PCT-US95-08429-9
                                                                                                                                                                                                                                                                                                                                                                                                                                                                              COMPUTER READABLE FORM:
                                                                                                                                                                                                                        708 ATDAPNGDLTPEM 720
                                                                                                                                                                          636 KREVKKGKVPEEV 648
                                                                                                                                                                                                                                                                                                                                                                                                                                                 NUMBER OF SEQUENCES:
                                                                                                                                                                                                                                                                                                                                                                                                                             INVENTION:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              amino acid
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   linear
                       621 DNEDEEEPAVEI -
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     FILING DATE: 1 CLASSIFICATION:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             CLASSIFICATION:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     FILING DATE:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               TELEPHONE:
                                                                                                                                                                                                                                                                                                                          PCT-US95-08429-9
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   TOPOLOGY:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  LENGTH:
                                                                                                                                                                                                                                                                                                                                                                                                   APPLICANT:
                                                                                                                                                                                                                                                                                                                                                                                                                             TITLE OF
                                                                                                                         QQ
                                                                                                                                                                            ó
                                                                                                                                                                                                                        QΩ
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            QC
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            g
                                                                          δŏ
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              δ
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             õ
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        38;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           -----LRNAKPP------250
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         292 YGIRFLRLSSEMFPFASHPVHGYKL-----APFASEVLAEAGRVAAELGHRLTTHPG 343
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           Q-----FTQLGSPRKEVVESAIRDLEYHDELLSLLKLPEQQN----RDAVMIIHM 389
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         404 EKPYECYICHARFTQSGTMKMHI-----LQKHTENVAKFHCPHCDTVIARKSDLGVHL 456
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            157 RKQHSYIEQGKKCRYCDAVFHERYALIQHQKSHKNEKRFKCDQCDYACRQERHMIMHKRT 516
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          ------PVCEE-LNIPMVLDYH----HHNICFDPAHLREGTLDISDPKLQERIANTWK 480
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              517 HTGEKPYACSHCDKTFRQKQLLDMHFKRYHDPNFVPAAF-----VCSKCGKTFT 565
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            178 LEQGELPPQEDPSWQKDPDYQPPAKKTKKTKKSKLRYTEEGKDVDVSVYDFEEEQQEGLL 237
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       347 HEKPF---KCSMCDYASVEVSKLKRHIRSHTGERPFQCSLCSYASRDTYKLKRHMRTHSG 403
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          GGQF-----GDKAATLER-FKRNYARL--SQSCKNRLVLENDDVGWTV----HDLL---- 433
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            525 EAKDKEQAVFELMRTFKLPGFEKINDMVPYDRDDENRPAPPVKAPKKKKGGKRK-RTTDE 583
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   97; Mismatches 252; Indels 304; Gaps
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 84 EGVKMAIEGLARMERRLQRATKRQKKQ--LEEDGIPVP-----SVVSRFPTAPYHHKS 134
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  58 VQMVMMEQLDPTLLQMKTEVMEGTVAPEAEAAVDDTQIITLQVVNMEEQPINIGELQLVQ 117
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       118 VPVPVTVPVATTSVEELQGAYENEVSKEGLAESEPMICHTLPLPEGFQVVKVGANGEVET 177
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  4 DAVEAIVEESETFIKGKERKTYQRRREGGQEEDACHLPQNQTDGGEVVQDV-----NSS 57
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      251 -------HLKNKPDKSKEPQDELGHKFVQELGLANARDIVKMLCWNEK
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   297 HLCGRAFRTVTLLRNHLNTHTGTRPHKCPD------CDMAFVTSGELVRHRRYKHT
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          RKGIKQKMHYSEPC------DGAVTPRDRRKHRPRVMTLPP------CPPDMDLMI
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  Ouery Match 3.3%; Score 115; DB 2; Length 727; Best Local Similarity 17.7%; Pred. No. 0.07;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   135 TNAEEREAKEPVL---KTHSKD--VEREAEIGVDD-----VVKME--
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   170 -----PAATNIIEPEDA---QDAAERGAARPPAVNSSYLPLP--
                                                                                                                                                                                            ATTORNEY/AGENT INFORMATION:
NAME: Parmelee, Steven W.
ERESISTATION NUMBER: 31,990
REFERENCE/DOCKET NUMBER: 14538A-11-1
                    APPLICATION NUMBER: US/08/475,844 FILING DATE: 07-JUN-1995 CLASSIFICATION: 536
                                                                                                       APPLICATION NUMBER: 08/261,680 FILING DATE: 17-JUN-1994 CLASSIFICATION: 536
                                                                                                                                                                                                                                                                                            TELECOMMUNICATION INFORMATION:
TELEPHONE: 206-467-9600
                                                                                                                                                                                                                                                                                                                                               TELEFAX: 415-543-5043
INFORMATION FOR SEQ ID NO: 9:
                                                                                                                                                                                                                                                                                                                                                                                              SEQUENCE CHARACTERISTICS:
LENGTH: 727 amino acids
TYPE: amino acid
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     Matches 140; Conservative
CURRENT APPLICATION DATA:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             ; MOLECULE TYPE: protein US-08-475-844-9
                                                                                            PRIOR APPLICATION DATA:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                              Linear
                                                                                                                                                                                                                                                                                                                                                                                                                                                                        TOPOLOGY:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       217
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  q
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  셤
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     à
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     q
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        ð
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     ö
```

346	343	403	389	456	433	516	480	565	524	620	583	662	635	707		
7 HLCGRAFRTVTLLRNHLNTHTGTRPHKCPDCDMAFVTSGELVRHRRYKHT			1 OFTQLGSPRKEVVESAIRDLEYHDELLSLLKLPEQQNRDAVMIIHM) GGQFGDKAATLER-FKRNYARLSQSCKNRLVLENDDVGWTVHDLL		PVCEE-LNIPMVLDYHHHNICFDPAHLREGTLDIS	7 HTGEKPYACSHCDKTFRQKQLLDMHFKRYHDPNFVPAAFVCSKCGKTFT		RRNTMARHADNCAGPDGVEGENGGETKKSKRGRKRKMRSKKEDSSDSENAEPDLD	5 EAKDKEQAVFELMRTFKLPGFEKINDMVPYDRDDENRPAPPVKAPKKKGGKRK-RTTDE				5 KREVKKGKVPEEV 648	708 ATDAPNGDLTPEM 720
297	292	347	344	404	390	457	434	517	481	266	525	621	584	663	636	307
qq	ΟŊ	qa	Qy	qq	Oy	qq	QŸ	QQ	Qy	QQ	QY	qq	Qy	qq	Qy	qq

Search completed: January 15, 2002, 13:51:13 Job time: 342 sec

•

```
January 15, 2002, 13:50:00; Search time 120.01 Seconds (without alignments) 404.900 Million cell updates/sec
                                                                                                                                                                                                                                                   1 MPSRKSKAAAALDIPQSESSI.....REVKKGKVPEEVEDEGEFDG 656
                                                                                                                                                                                                                                                                                                                                                                                                                 522463
GenCore version 4.5
Copyright (c) 1993 - 2000 Compugen Ltd.
                                                                                                                                                                                                                                                                                                                                                                                                 Total number of hits satisfying chosen parameters:
                                                                                                                                                                                                                                                                                                                                                              522463 seqs, 74073290 residues
                                                                                  OM protein - protein search, using sw model
                                                                                                                                                                                                                                                                                                  BLOSUM62
Gapop 10.0 , Gapext 0.5
                                                                                                                                                                                                                                                                                                                                                                                                                                              Minimum DB seq length: 0
Maximum DB seq length: 2000000000
                                                                                                                                                                                                               US-09-724-296-36
                                                                                                                                                                                                                                   Perfect score:
                                                                                                                                                                                                                                                                                               Scoring table:
                                                                                                                                                                                                                                                       Sequence:
                                                                                                                                                                                                                                                                                                                                                                Searched:
                                                                                                                          Run on:
```

Post-processing: Minimum Match 0% Maximum Match 100% Listing first 45 summaries A_Geneseq_1101:* Database

/SIDSB/gcgdata/genesseq/geneseepp/AA1991.DAT:*
/SIDSB/gcgdata/genesseq/geneseepp/AA1992.DAT:*
/SIDSB/gcgdata/genesseq/geneseepp/AA1993.DAT:*
/SIDSB/gcgdata/genesseq/geneseepp/AA1995.DAT:*
/SIDSB/gcgdata/geneseeq/geneseepp/AA1995.DAT:*
/SIDSB/gcgdata/geneseeq/geneseepp/AA1995.DAT:*
/SIDSB/gcgdata/geneseeq/geneseepp/AA1995.DAT:*
/SIDSB/gcgdata/geneseeq/geneseepp/AA1995.DAT:* /SIDS8/gcgdata/geneseq/geneseqp/AA1990.DAT:* /SIDS8/gcgdata/geneseq/geneseqp/AA2000.DAT:*/SIDS8/gcgdata/geneseqf/geneseqp/AA2001.DAT:* /SIDSB/gcgdata/geneseq/geneseqp/AA1983.DAT:*
/SIDSB/gcgdata/geneseq/geneseqp/AA1984.DAT:*
/SIDSB/gcgdata/geneseq/geneseqp/AA1986.DAT:*
/SIDSB/gcgdata/geneseq/geneseqp/AA1986.DAT:*
/SIDSB/gcgdata/geneseq/geneseqp/AA1986.DAT:*
/SIDSB/gcgdata/geneseq/geneseqp/AA1989.DAT:*
/SIDSB/gcgdata/geneseq/geneseqp/AA1988.DAT:* /gcgdata/geneseq/geneseqp/AA1989.DAT:* /SIDS8/gcgdata/geneseq/genesegp/AA1980.DAT:*/SIDS8/gcgdata/geneseq/genesegp/AA1981.DAT:* /SIDS8/gcgdata/geneseq/geneseqp/AA1998.DAT: /SIDS8/gcgdata/geneseq/geneseqp/AA1999.DAT: /SIDS8/gcgdata/geneseq/geneseqp/AA1982. 110... 111... 112... 115... 116... 118... 118... 118... 118...

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

:	Description	account encountries	N Crassa deltable	B subtilie deltas	GST signal pontido	CST signal peptide	S nombe dollarone.	D rediodurand dol	Human Aciens 1 220	Human protoin com	A human process sequ	Human foetal liver
SUMMARIES		7 AAW01618	L AAY44502	L AAY44503	AAY44498	AAY44500	AAY44499	AAY44505	AAY85657	AAB94420	AAB19002	AAB29444
% Query Watch Length DD	ייייי ביייייי	656 17	626 21	626 21	828 21	600 21	371 21	294 21	1341 21	914 22	747 21	747 21
% Query Match		99.7	94.5	94.5	29.0	27.7	27.6	7.5	4.6	4.1	4.0	4.0
or COS		3472	3292	3292	1011.5	963	096	262.5	160.5	143.5	139	139
Result		1	7	Э	4	S	9	7	8	0	10	11

Disclosure; Page 10-13; 18pp; Japanese.

Human cell cycle p Peptide #967 encod Peptide #987 encod Peptide #941 encod Mouse Rimc, a nove	Polyl Polyl Polyl Polyl Acinu	acid seq umour sup de #2892 de #2869 estin. R estin. pro	HURCR protein seque HURCR protein Segue Murine protein, SE Murine protein, SE Novel protein kina Murine protein, SE C albicans apoptos Plasmodium falcipa Human metastasisa	
AAY50976 AAM4533 AAM26950 AAB73488 AAB73488	30033 30033 30033 4088	2777777777	1245 3488 3488 3488 3488 3488 3488 3488 34	378 378 378 378 361 361 365 386
22 22 22 22 22 22	22 22 21 21 21 21	22 22 22 13 15	22 22 22 22 23 21 23 18	20 21 21 21 21 21 19
	98861	800 801	2453 1272 1272 1490 2243 1072 1382	7000033
		200	127.5 126.5 126.5 126.5 126.5 125.5	
12 13 14 15 16		225 25 26 28 29	30 33 33 34 34 37	8 0 0 1 2 8 4 4 4 4 4 4 4 4 4 4 4 4 4 4 4 4 4 4

ALIGNMENTS

```
DNA repair enzyme; cyclobutane-type dimer; (6-4)-bound product; exposure; UV light; ultraviolet; suntan cream; prevention; treatment; skin cancer.
                                                                                                                                                                                                                                                                                             DNA repair enzyme and related DNA - specifically recognises cyclobutane-type dimer and (6-4)-bound product induced by exposure
                                                                       Neurospora crassa DNA repair enzyme.
      AAW01618 standard; Protein; 656 AA.
                                                                                                                                                                                                                       95JP-0094137.
                                                                                                                                                                                                   95JP-0094137.
                                                   20-AUG-1997 (first entry)
                                                                                                                                                                                                                                          (SHIS ) SHISEIDO CO LID
                                                                                                                                                                                                                                                              WPI; 1996-512663/51.
                                                                                                                                     Neurospora crassa.
                                                                                                                                                                                                                                                                         N-PSDB; AAT58286
                                                                                                                                                         JP08266275-A.
                                                                                                                                                                                                 29-MAR-1995;
                                                                                                                                                                                                                    29-MAR-1995;
                                                                                                                                                                              15-OCT-1996
                                                                                                                                                                                                                                                                                                                   to UV light
                            AAWOT6T87
AAW01618
```

```
228
                                                                                                                                                                                                                                ;
0
   The present sequence shows a DNA repair enzyme which specifically recognises a cyclobutane-type dimer and (6-4)-bound product induced, respectively from Tr and TC sequences in DNA, by exposure to UV light. The enzyme cleaves these sequences at the 5'-side. The DNA repair enzyme can be used in, e.g. suntan cream for prevention and treatment of skin cancer. Conventional endonuclease acts on either the cyclobutane-type dimer or the (6-4)-bound product in damaged DNA, while the enzyme of the invention can act on both of them.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                       421 enddvgwtvhdllpvceelnipmvldyhhnicfdpahlregtldisdpklqeriantwk 480
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              481 RKGIKQKMHYSEPCDGAVTPRDRRKHRPRVMTLPPCPPDMDLMIEAKDKEQAVFELMRTF 540
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            541 KLPGFEKINDMVPYDRDDENRPAPPVKAPKKKKGGKRKRTTDEEAAEPEEVDTAADDVKD 600
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          AQDAAERGAARPPAVNSSYLPLPWKGRLGYACLNTYLRNAKPPIFSSRTCRMASIVDHRH 240
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         ENDDVGWTVHDLLPVCEELNIPMVLDYHHHNICFDPAHLREGTLDISDPKLQERIANTWK 480
                                                                                                                                                                                                                                                                                                                                             PLQFEDEPEHHLKNKPDKSKEPQDELGHKFVQELGLANARDIVKMLCWNEKYGIRFLRLS 300
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   SEMPPFASHPVHGYKLAPFASEVLAEAGRVAAELGHRLTTHPGQFTQLGSPRKEVVESAI 360
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 RDLEYHDELLSLLKLPEQQNRDAVMIIHMGGQFGDKAATLERFKRNYARLSQSCKNRLVL 420
                                                                                                                                                                                                                                                                                                                             ELAGRAMGKDANGHCLREGKEQEEGVKMAIEGLARMERRLQRATKRQKKQLEEDGIPVPS 120
                                                                                                                                                                                                                                                                                                                                                                                           VVSRFPTAPYHHKSTNAEEREAKEPVLKTHSKDVEREAEIGVDDVVKMEPAATNIIEPED 180
                                                                                                                                                                                                                                                                                 Gaps
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     Delta228-UVDE; ultraviolet damage endonuclease; GST signal peptide;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                1 MPSRKSKAAALDTPQSESSTFSSTLDSSAPSPARNLRRSGRNILQPSSEKDRDHEKRSGE
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           601 APEGPKEVPEEERAMGGPYNRVYWPLGCEEWLKPKKREVKKGKVPEEVEDEGEFDG 656
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            0;
                                                                                                                                                                                                   DB 17; Length 656;
                                                                                                                                                                                                                                   Indels
                                                                                                                                                                                                   99.7%; Score 3472; DB 17;
99.7%; Pred. No. 1.5e-295;
Live 1; Mismatches 1;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       crassa delta228-UV damage endonuclease.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          AAY44502 standard; Protein; 626 AA.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          27-MAR-2000 (first entry)
                                                                                                                                                                                                                                   Matches 654; Conservative
                                                                                                                                                                                                                      Best Local Similarity
                                                                                                                                                       656 AA;
                                                                                                                                                         Sequence
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            AAY44502;
                                                                                                                                                                                                      Query Match
                                                                                                                                                                                                                                                                                                                                                                                                                                                           181
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        241
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      301
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     421
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        301
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           AAY44502
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 OD
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 Dp
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 δŏ
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                ŏ
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      qq
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     ōλ
                                                                                                                                                                                                                                                                                                                                                                                                                           qq
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        ó
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        ð
x00000000000x
                                                                                                                                                                                                                                                                     ŏ
                                                                                                                                                                                                                                                                                                  g
                                                                                                                                                                                                                                                                                                                                οy
                                                                                                                                                                                                                                                                                                                                                              d
                                                                                                                                                                                                                                                                                                                                                                                             Qγ
                                                                                                                                                                                                                                                                                                                                                                                                                                                           Q
```

```
The present sequence is delta228-UV damage endonuclease from N. crassa. Delta228-UVDE is a 288 amino acid deletion of the N-terminal of the Uvel+ gene product. This is expressed in frame with a GST leader sequence to generate a fusion protein. This provides stable endonuclease fragments for cleaving a double-stranded DNA molecule that has a distorted structure resulting from UV radiation damage, a photoproduct, an abasic site, mismatched nucleotide pairing, aplatinum diaduct, an intercalated molecule or alkylation of a nucleotide. Uvelp can be used in compositions for internal or topical application and as a therapeutic agent for skin
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                361 RDLEYHDELLSLLKLPEQQNRDAVMIIHMGGQFGDKAATLERFKRNYARLSQSCKNRLVL 420
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  181 aqdaaergaarppavnssylplpwkgrlgyaclntylrnakppifssrtcrmasivdhrh 240
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                SEMFPFASHPVHGYKLAPFASEVLAEAGRVAAELGHRLTTHPGQFTQLGSPRKEVVESAI 360
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   181 AQDAAERGAARPPAVNSSYLPLPWKGRLGYACLNTYLRNAKPPIFSSRTCRMASIVDHRH 240
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                61 ELAGRAMGKDANGHCLREGKEQEEGVKMAIEGLARMERRLQRATKRQKKQLEEDGIPVPS 120
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  121 VVSRFPTAPYHHKSTNAEEREAKEPVLKTHSKDVEREAEIGVDDVVKMEPAATNIIEPED 180
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               PLQFEDEPEHHLKNKPDKSKEPQDELGHKFVQELGLANARDIVKMLCWNEKYGIRFLRLS 300
                                                                                                                                                                                                                                                                                                                                                                 A new truncated ultraviolet damage endonuclease for treatment of skin
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                30; Gaps
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               1 MPSRKSKAAALDTPQSESSTFSSTLDSSAPSPARNLRRSGRNILQPSSEKDRDHEKRSGE 60
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               1 mpsrkskaaaldtpgsesstfsstldssapsparnlrrsgrnilgpssekdrdhekrsge
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             plqfedepehhlknkpdkskepqdelghkfvqelglanardivkmlcw------
qlutathione-S-transferase signal peptide; uvel+ gene product; UV irradiation; DNA damage; UV radiation damage; photoproduct; abasic site; aplatinum diaduct; mismatched nucleotide pairing;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               Length 626;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  Indels
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 Score 3292; DB 21;
Pred. No. 8.5e-280;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  0; Mismatches
                                                      nucleotide alkylation; skin cancer.
                                                                                                                                                                                                                                                                                                                                                                                                                    Claim 16; Page 59; 133pp; English.
                                                                                                                                                                                                                                                                                                      Avery AM;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  94.5%;
95.4%;
                                                                                                                                                                                     99WO-US12910
                                                                                                                                                                                                                                    99US-0134752
                                                                                                                                                                                                                      98US-0088521
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   Conservative
                                                                                                                                                                                                                                                                                                      Kaur B,
                                                                                                                                                                                                                                                                                                                                  WPI; 2000-116417/10.
                                                                                                                                                                                                                                                                     EMORY.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    Local Similarity
                                                                                     Neurospora crassa.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     626 AA;
                                                                                                                                                                                                                                                                     (UYEM-) UNIV
                                                                                                                      W09963828-A1
                                                                                                                                                                                     08-JUN-1999;
                                                                                                                                                                                                                                    18-MAY-1999;
                                                                                                                                                                                                                                                                                                      Doetsch PW,
                                                                                                                                                       16-DEC-1999
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   626;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     Sequence
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      Query Match
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       cancers
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     Matches
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    61
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      241
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      241
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       301
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    Dp
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      QO
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    QQ
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       qq
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       ŏ
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   qq
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      òγ
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    q
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     δý
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      δŏ
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     δ
```

```
The present sequence is delta228-UV damage endonuclease from B. subtilis. Delta228-UVDE is a 288 amino acid deletion of the N-terminal of the uvel+ gene product. This is expressed in frame with a GST leader sequence to generate a fusion protein. This provides stable endonuclease fragments for cleaving a double-stranded DNA molecule that has a distorted structure resulting from UV radiation damage, a photoproduct, an abasic site, mismatched nucleotide pairing, aplatinum diaduct, an intercalated molecule or alkylation of a nucleotide. Uvelp can be used in compositions for internal or topical application and as a therapeutic agent for skin
                                    465
                                                                                                                          541 KLPGFEKINDMVPYDRDDENRPAPPVKAPKKKKGGKRKRTTDEEAAEPEEVDTAADDVKD 600
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        new truncated ultraviolet damage endonuclease for treatment of skin
                                                                                                                                                                                                                                                                                                                                                                                                                         Delta228-UVDE; ultraviolet damage endonuclease; GST signal peptide; glutathione-S-transferase signal peptide; uvel+ gene product; UV irradiation; DNA damage; UV radiation damage; photoproduct; abasic site; aplatinum diaduct; mismatched nucleotide pairing; nucleotide alkylation; skin cancer.
ENDDVGWTVHDLLPVCEELNIPMVLDYHHHNICFDPAHLREGTLDISDPKLQERIANTWK
                  RKGIKQKMHYSEPCDGAVTPRDRRKHRPRVMTLPPCPPDMDLMIEAKDKEQAVFELMRTF
                                                                              577 ----evpeeeramggpynrvywplgceewlkpkkrevkkgkvpeevedegefdg 626
                                                                                                                                                                                       601 APEGPKEVPEEERAMGGPYNRVYWPLGCEEWLKPKKREVKKGKVPEEVEDEGEFDG
                                                                                                                                                                                                                                                                                                                                                                                             B. subtilis delta228-UV damage endonuclease.
                                                                                                                                                                                                                                                                                                  AAY44503 standard; Protein; 626 AA.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      Claim 16; Page 59; 133pp; English.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                Avery AM,
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    99WO-US12910
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                99US-0134752
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 98US-0088521
                                                                                                                                                                                                                                                                                                                                                                 (first entry)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             Kaur B,
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          WPI; 2000-116417/10
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              (UYEM-) UNIV EMORY.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       Bacillus subtilis.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              626 AA;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    W09963828-A1
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  08-JUN-1999;
                                                                                                                                                                                                                                                                                                                                                                27-MAR-2000
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                18-MAY-1999;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 08-JUN-1998;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             Doetsch PW,
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     16-DEC-1999
                                                                                                                                                                                                                                                                                                                                  AAY44503
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               Seguence
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           cancers
 421
                               406
                                                                                           466
                                                              181
                               a
                                                              ð
                                                                                          g
                                                                                                                         ŏ
                                                                                                                                                        g
                                                                                                                                                                                       ò
                                                                                                                                                                                                                   g
```

```
ELAGRMMGKDANGHCLREGKEQEEGVKMAIEGLARMERRLQRATKRQKKQLEEDGIPVPS 120
                                                               VVSRFPTAPYHHKSTNAEEREAKEPVLKTHSKDVEREAEIGVDDVVKMEPAATNIIEPED 180
                                                                                                                AQDAAERGAARPPAVNSSYLPLPWKGRLGYACLNTYLRNAKPPIFSSRTCRMASIVDHRH 240
                                                                                                                                                                                 181 aqdaaergaarppavnssylplpwkgrlgyacintylrnakppifssrtcrmasivdhrh 240
                                                                                                                                                                                                           PLQFEDEPEHHLKNKPDKSKEPQDELGHKFVQELGLANARDIVKMLCWNEKYGIRFLRLS 300
                                                                                                                                                                                                                       SEMFPFASHPVHGYKLAPFASEVLAEAGRVAAELGHRLTTHPGQFTQLGSPRKEVVESAI 360
                                                                                                                                                                                                                                                                                        ---fpfashpvhgyklapfasevlaeagrvaaelghrltthpgqftqlgsprkevvesai 345
                                                                                                                                                                                                                                                                                                                                ENDDVGWTVHDLLPVCEELNIPMVLDYHHHNICFDPAHLREGTLDISDPKLQERIANTWK 480
                                                                                                                                                                                                                                                                                                                                                                                   576
                                                                                                                                                                                                                                                                                                                 361 RDLEYHDELLSLLKLPEQQNRDAVMIIHMGGQFGDKAATLERFKRNYARLSQSCKNRLVL 420
                                                                                                                                                                                                                                                                                                                                                                                                                        RKGIKQKMHYSEPCDGAVTPRDRRKHRPRVMTLPPCPPDMDLMIEAKDKEQAVFELMRTF 540
                                                                                                                                                                                                                                                                                                                                                                                                                                                                           KLPGFEKINDMVPYDRDDENRPAPPVKAPKKKGGKRKRTTDEEAAEPEEVDTAADDVKD 600
            MPSRKSKAAAALDTPQSESSTFSSTLDSSAPSPARNLRRSGRNILQPSSEKDRDHEKRSGE
                                                                                                                                                                                                                                                                                                                                                                                                                                    601 APEGPKEVPEEERAMGGPYNRVYWPLGCEEWLKPKKREVKKGKVPEEVEDEGEFDG 656
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          GST signal peptide; ultraviolet damage endonuclease; UVDE; Uvelp; UV irradiation; DNA damage; UV radiation damage; fusion protein; skin cancer; glutathione-S-transferase.
                                                                                                                                                                     signal peptide and S. pombe UVDE fusion protein.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    1..229
/label= GST_signal_peptide
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   /note= "Encoded by aaaa"
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             'note= "Encoded by gtt"
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      /note= "Encoded by gtt"
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         /note= "S. pombe UVDE"
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        Location/Qualifiers
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      AAY44498 standard; Protein; 828
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         27-MAR-2000 (first entry)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                Schizosaccharomyces pombe
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             230..828
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          Misc-difference 147
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     Misc-difference 11
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                Misc-difference 85
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             Synthetic.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                AAY44498
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   Peptide
                                                 61
                                                                                                                                                                                                                                                              301
                                                                                                                                                                                                                                                                                                                                                                                                                                                  466
                                                                                                                                                                                                                                                                                                                                                                                              406
                                                                                                                                                                                                                                                                                                                                                                                                                        481
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            RESULT
AAY44498
                    qq
                                                ŏ
                                                                          Op
                                                                                                    ó
                                                                                                                          q
                                                                                                                                                       Óγ
                                                                                                                                                                                g
                                                                                                                                                                                                          ò
                                                                                                                                                                                                                                  qq
                                                                                                                                                                                                                                                              Qγ
                                                                                                                                                                                                                                                                                        Op
D
                                                                                                                                                                                                                                                                                                               ŏ
                                                                                                                                                                                                                                                                                                                                          QΩ
                                                                                                                                                                                                                                                                                                                                                                    δ
                                                                                                                                                                                                                                                                                                                                                                                            QQ
                                                                                                                                                                                                                                                                                                                                                                                                                       ŏ
                                                                                                                                                                                                                                                                                                                                                                                                                                               QQ
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       G.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                qq
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                           δ
```

7

Gaps

30;

. 0

Length 626; Indels

94.5%; Score 3292; DB 21; 95.4%; Pred. No. 8.5e-280;

0; Mismatches

Best Local Similarity 95.4 Matches 626; Conservative

Query Match

```
18;
                                                                                                                                                                                                                                                                                                                                                                                                        The present sequence is a fusion protein comprising the GST signal peptide and full length ultraviolet damage endonuclease (UVDE) from Schizoascacharonyces pombe. UVDE genes are isolated by introducing a foreign cDNA library into a repair-deficient E. coli strain and selecting for complemented cells by UV irradiation of the transformants. This provides stable endonuclease fragments for cleaving a double-stranded DNA molecule that has a distorted structure resulting from UV radiation damage, a photoproduct, an abasic site, mismatched nucleotide pairing, aplatinm diaduct, an intercalated molecule or alkylation of a nucleotide. Uvelp can be used in compositions for internal or topical application and as a therapeutic agent for skin cancers.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               | : |::|| | | ::|::|| :|
-----qrdglesvkqlgtqnvldliklvewnhnfg 536
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           414 CKNRLVLENDDVGWTVHDLLPVCEELNIPMVLDYHHHNICFDPAHLREGTLDISDPKLQE 473
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             -----seegiskatkkvkqkeeeeyveevdek 413
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               174 NIIEPEDAQDAAERGAARPPAVNSSYLPLPWKGRLGYACLNTYLRNAKPPIFSSRTCRMA 233
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      234 SIVDHRHPLQFEDEPEHHLKNKPDKSKEPQDELGHKFVQELGLANARDIVKMLCWNEKYG 293
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   294 IRFLRLSSEMFPFASHPVHGYKLAPFASEVLAEAGRVAAELGHRLTTHPGQFTQLGSPRK 353
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            354 EVVESAIRDLEYHDELLSLLKLPEQQNRDAVMIIHMGGQFGDKAATLERFKRNYARLSQS 413
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   339 arsskkkatdsvsdkidesvasydsst----hlrrssr----skkpvnynsssese- 386
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            63 AGRMMGKDANGHCLREGKEQEEGVKMAIEGLARMERRLQRATKRQKKQLEED----- 114
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      115 GIPVPSVVSRF-PTAPYHHKSTNAEEREAKEPVLKTHSKDVEREAEIGVDDVVKMEPAAT 173
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           414 slknesssdefepvvpeqletpiskrrrs----rssaknlekestmnlddha----- 461
                                                                                                                                                                                                                                                                                                                     new truncated ultraviolet damage endonuclease for treatment of skin
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    29.0%; Score 1011.5; DB 21; Length 828; 36.6%; Pred. No. 1.1e-79; tive 88; Mismatches 149; Indels 181; Gaps
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 3 SRKSKAAALDTPQSESSTFSSTLDSSAPSPARNLRRSGRNILQPSSEKDRDHEKRSGEEL 62
                                                                                                                                                                                                                                                                                                                                                                                Claim 7; Page 51; 133pp; English.
                                                                                                                                                                                                                              Avery AM;
                                                                                         99WO-US12910.
                                                                                                                               98US-0088521.
                                                                                                                                                 99US-0134752
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                Conservative
                                                                                                                                                                                                                            Kaur B,
                                                                                                                                                                                                                                                                 WPI; 2000-116417/10.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          Best Local Similarity
                                                                                                                                                                                      (UYEM-) UNIV EMORY
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   828 AA;
                                                                                                                                                                                                                                                                                   N-PSDB; AAZ29858
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   505 ti------
                W09963828-A1
                                                                                         08-JUN-1999;
                                                                                                                               08-JUN-1998;
                                                                                                                                                   18-MAY-1999;
                                                                                                                                                                                                                            Doetsch PW,
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                Matches 241;
                                                      16-DEC-1999
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     Sequence
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          Query Match
                                                                                                                                                                                                                                                                                                                                              cancers
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      462
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    387
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          a
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            ŏ
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       ŏ
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                δ
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                δ
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   ò
```

```
The present sequence is a fusion protein comprising the GST signal peptide and delta228 truncated ultraviolet damage endonuclease (UVDE) from S. pombe. Delta228-UVDE is a 288 amino acid deletaion of the N-terminal of the S. pombe uvel+ gene product. This is expressed in frame with a GST leader sequence. Stable endonuclease fragments can be produced for cleaving a double-stranded DNA molecule that has a distorted structure resulting from UV radiation damage, a photoproduct, an abasic site, mismatched nucleotide pairing, aplatinum diaduct, an intercalated molecule or alkylation of a nucleotide. Uvelp can be used in compositions for internal or topical application and as a therapeutic agent for skin
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    A new truncated ultraviolet damage endonuclease for treatment of skin
                      712 tiretwtrkgitgkghysesadptaisgmkrrahsdrvfdippcdptmdlmieakekega 771
                                                                     533 VFELMRTFKLPGFEKINDMVPYDRDDENRPAPPVKAPKKKKGGKRKRTTDEEAAEPEEVD 592
474 RIANTWKRKGIKQKMHYSEPCD-GAVTPRDRRKHRPRVMTLPPCPPDMDLMIEAKDKEQA 532
                                                                                                       -----qnppcp----- 794
                                                                                                                                         593 TAADDVKDAPEGPKEVPEEERAMGGPYNRVYWPLGCEEWLKPKKREVKKGKVPEEVEDE 651
                                                                                                                                                                         -----yyppgaekrltarkrrsrk----eeveed 826
                                                                                                                                                                                                                                                                                                                                                                                                          GST signal peptide; glutathione-S-transferase signal peptide; ultraviolet damage endonuclease; UVDE; delta228-UVDE; repair-deficient E. coli strain; UV irradiation; DNA damage; UV radiation damage; photopproduct; abasic site; aplatinum diaduct; mismatched nucleotide pairing; nucleotide alkylation;
                                                                                                                                                                                                                                                                                                                                                                        CST signal peptide and delta228 S. pombe UVDE fusion protein.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            /note= "Encoded by aaaa"
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                /note= "Encoded by gtt"
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                /note= "Encoded by gtt"
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           Disclosure; Page 56; 133pp; English.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           Location/Qualifiers
                                                                                                                                                                                                                                                                     AAY44500 standard; Protein; 600 AA.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  fusion protein; skin cancer.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         99WO-US12910.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           98US-0088521.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            99US-0134752.
                                                                                                                                                                                                                                                                                                                                            27-MAR-2000 (first entry)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        Schizosaccharomyces pombe.
                                                                                                           772 vfelcrryel-----
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   Kaur B,
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   WPI; 2000-116417/10.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                UYEM-) UNIV EMORY.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            Misc-difference 11
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             Misc-difference 85
                                                                                                                                                                                795 --ydqtrdg----
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      N-PSDB; AAZ29860
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 Misc-difference
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      W09963828-A1.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         08-JUN-1999;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           08-JUN-1998;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              18-MAY-1999;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   Doetsch PW,
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       16-DEC-1999
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             Synthetic
                                                                                                                                                                                                                                                                                                        AAY44500;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         cancers.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           cancers
                                                                                                                                                                                                                                                      AAY44500
                                                                                                                                                                                                                                     RESULT
                                                                                                                                                                                                                                                                                                        Сp
                                                                         Ωy
                                                                                                                                                ò
     ò
```

```
cancers
                                                                                                                                                                                                                                                                                                                                                                                                                                              cancers
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           106
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          AAY44505
      οy
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 qq
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      pp
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               δλ
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           qq
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      Óγ
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             QQ
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    qq
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         Ωp
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            Qγ
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                δλ
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    δ
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            δò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   Q
                                                                                                                                  13;
                                                                                                                                Gaps
                                                                                                                                                                    145 PVLKTHSKDVEREAEIGVDDVVKMEPAATNIIEPEDAQDAAERGAARPPAVNSSYL---- 200
                                                                                                                                                                                                           176 pklvcfkkrie----aipqidkylksskyiawplqgwqatfgggdhpp--ksdhlvprg 228
                                                                                                                                                                                                                                                            -----PLPWKGRLGYACLNTYLRNAKPPIFSSRTCRMASIVDHRHPLQFE 245
                                                                                                                                                                                                                                                                                                                                              246 DEPEHHLKNKPDKSKEPQDELGHKFVQELGLANARDIVKMLCWNEKYGIRFLRLSSEMFP 305
                                                                                                                                                                                                                                                                                                                                                                                                                                       365
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            425
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            426 GWTVHDLLPVCEELNIPMVLDYHHHNICFDPAHLREGTLDISDPKLQERIANTWKRKGIK 485
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   554 ----ydqtrdg--- 573
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    545 FEKINDMVPYDRDDENRPAPPVKAPKKKGGKRKRTTDEEAAEPEEVDTAADDVKDAPEG 604
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       Delta228-UVDE; ultraviolet damage endonuclease; GST signal peptide;
                                                                                                                                                                                                                                                                                     | : |::|| | |::|::|| | |::|::|| |::|| |::|| |::|| |::|| |::|| |::|| |::|| |::|| |::|| |::|| |::|| |::|| |::|| |::|| |::|| |::|| |::|| |::|| |::|| |::|| |::|| |::|| |::|| |::|| |::|| |::|| |::|| |::|| |::|| |::|| |::|| |::|| |::|| |::|| |::|| |::|| |::|| |::|| |::|| |::|| |::|| |::|| |::|| |::|| |::|| |::|| |::|| |::|| |::|| |::|| |::|| |::|| |::|| |::|| |::|| |::|| |::|| |::|| |::|| |::|| |::|| |::|| |::|| |::|| |::|| |::|| |::|| |::|| |::|| |::|| |::|| |::|| |::|| |::|| |::|| |::|| |::|| |::|| |::|| |::|| |::|| |::|| |::|| |::|| |::|| |::|| |::|| |::|| |::|| |::|| |::|| |::|| |::|| |::|| |::|| |::|| |::|| |::|| |::|| |::|| |::|| |::|| |::|| |::|| |::|| |::|| |::|| |::|| |::|| |::|| |::|| |::|| |::|| |::|| |::|| |::|| |::|| |::|| |::|| |::|| |::|| |::|| |::|| |::|| |::|| |::|| |::|| |::|| |::|| |::|| |::|| |::|| |::|| |::|| |::|| |::|| |::|| |::|| |::|| |::|| |::|| |::|| |::|| |::|| |::|| |::|| |::|| |::|| |::|| |::|| |::|| |::|| |::|| |::|| |::|| |::|| |::|| |::|| |::|| |::|| |::|| |::|| |::|| |::|| |::|| |::|| |::|| |::|| |::|| |::|| |::|| |::|| |::|| |::|| |::|| |::|| |::|| |::|| |::|| |::|| |::|| |::|| |::|| |::|| |::|| |::|| |::|| |::|| |::|| |::|| |::|| |::|| |::|| |::|| |::|| |::|| |::|| |::|| |::|| |::|| |::|| |::|| |::|| |::|| |::|| |::|| |::|| |::|| |::|| |::|| |::|| |::|| |::|| |::|| |::|| |::|| |::|| |::|| |::|| |::|| |::|| |::|| |::|| |::|| |::|| |::|| |::|| |::|| |::|| |::|| |::|| |::|| |::|| |::|| |::|| |::|| |::|| |::|| |::|| |::|| |::|| |::|| |::|| |::|| |::|| |::|| |::|| |::|| |::|| |::|| |::|| |::|| |::|| |::|| |::|| |::|| |::|| |::|| |::|| |::|| |::|| |::|| |::|| |::|| |::|| |::|| |::|| |::|| |::|| |::|| |::|| |::|| |::|| |::|| |::|| |::|| |::|| |::|| |::|| |::|| |::|| |::|| |::|| |::|| |::|| |::|| |::|| |::|| |::|| |::|| |::|| |::|| |::|| |::|| |::|| |::|| |::|| |::|| |::|| |::|| |::|| |::|| |::|| |::|| |::|| |::|| |::|| |::|| |::|| |::|| |::|| |::|| |::|| |::|| |::|| |::|| |::|| |::|| |::|| |::|| |::|| |::|| |::|| |::|| |::|| |:|| |::|| |::|| |::|| |::|| |::|| |::|| |::|| |::|| |::|| |::||
                                                                                                                                                                                                                                                                                                                                                                                                                                   FASHPVHGYKLAPFASEVLAEAGRVAAELGHRLTTHPGQFTQLGSPRKEVVESAIRDLEY
                                                                                                                                                                                                                                                                                                                                                                                                                                                          HDELLSLLKLPEQQNRDAVMIIHMGGQFGDKAATLERFKRNYARLSQSCKNRLVLENDDV
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 QKMHYSEPCD-GAVTPRDRRKHRPRVMTLPPCPPDMDLMIEAKDKEQAVFELMRTFKLPG
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        60; Mismatches 125; Indels 124;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           glutathione-S-transferase signal peptide; uvel+ gene product; UV irradiation; DNA damage; UV radiation damage; photoproduct; abasic site; aplatinum diaduct; mismatched nucleotide pairing;
                                                                                       Length 600;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        605 PKEVPEEERAMGGPYNRVYWPLGCEEWLKPKKREVKKGKVPEEVEDE 651
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    27.7%; Score 963; DB 21; 41.4%; Pred. No. 1.2e-75;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              pombe delta228-UV damage endonuclease
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              nucleotide alkylation; skin cancer.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             AAY44499 standard; Protein; 371
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          99WO-US12910
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     98US-0088521
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            99US-0134752
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    (first entry)
                                                                                                                              Conservative
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       Schizosaccharomyces pombe
                                                                                                       Local Similarity
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     (UYEM-) UNIV EMORY
                600 AA;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  27-MAR-2000
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       08-JUN-1999;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            18-MAY-1999;
                                                                                                                       Matches 218;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   08-JUN-1998
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              16-DEC-1999
                  Sequence
                                                                                  Query Match
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    9
                                                                                                                                                                                                                                                         201
                                                                                                                                                                                                                                                                                                                                                                                          279
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 380
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        366
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               486
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         AAY44499
χg
                                                                                                                                                                       õ
                                                                                                                                                                                                                                                         ò
                                                                                                                                                                                                                                                                                                q
                                                                                                                                                                                                                                                                                                                                              ò
                                                                                                                                                                                                                                                                                                                                                                                     셤
                                                                                                                                                                                                                                                                                                                                                                                                                                   ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                       g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 a
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      셤
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             0
```

```
The present sequence is delta228-UV damage endonuclease. Delta228-UVDE is a 288 amino acid deletion of the N-terminal of the S. pombe uvel+ gene product. This is expressed in frame with a GST leader sequence to generate a fusion protein. This provides stable endonuclease fragments for cleaving a double-stranded DNA molecule that has a distorted structure resulting from UV radiation damage, a photoproduct, an abasic site, mismatched nucleotide pairing, aplatinum diaduct, an intercalated molecule or alkylation of a nucleotide. Uvelp can be used in compositions for internal or topical application and as a therapeutic agent for skin
                                                                                                                                                          new truncated ultraviolet damage endonuclease for treatment of skin
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 | : |::|| | |::|::|| :|| |:||---qrdglesvkqlgtqnvldliklvewnhnfgihfmrvssdlfpfashakygytl-efa 105
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   201 PLPWKGRLGYACLNTYLRNAKPPIFSSRTCRMASIVDHRHPLQFEDEPEHHLKNKPDKSK 260
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               261 EPQDELGHKFVQELGLANARDIVKMLCWNEKYGIRFLRLSSEMFPFASHPVHGYKLAPFA 320
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 321 SEVLAEAGRVAAELGHRLTTHPGQFTQLGSPRKEVVESAIRDLEYHDELLSLLKLPEQQN 380
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             381 RDAVMIIHMGGQFGDKAATLERFKRNYARLSQSCKNRLVLENDDVGWTVHDLLPVCEELN 440
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           441 IPMVLDYHHHNICFDPAHLREGTLDISDPKLQERIANTWKRKGIKQKMHYSEPCD-GAVT 499
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         560 NRPAPPVKAPKKKKGGKRKRTTDEEAAEPEEVDTAADDVKDAPEGPKEVPEEERAMGGPY 619
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          500 PRDRRKHRPRVMTLPPCPPDMDLMIEAKDKEQAVFELMRTFKLPGFEKINDMVPYDRDDE 559
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           Delta228-UVDE; ultraviolet damage endonuclease; GST signal peptide;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            226 iplvldwhhhniv--pgtlregsldlm--pliptiretwtrkgitgkghysesadptais
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         98;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 27.6%; Score 960; DB 21; Length 371; 45.4%; Pred. No. 1.1e-75;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         Indels
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 nppcp----ydqtrdg----
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            D. radiodurans delta228-UV damage endonuclease
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      Mismatches
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           620 NRVYWPLGCEEWLKPKKREVKKGKVPEEVEDE 651
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            ---yyppgaekrltarkrrsrk---eeveed 369
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             AA
                                                                                                                                                                                                                                                   Claim 13; Page 53; 133pp; English.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      54;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          AAY44505 standard; Protein; 294
   Avery AM;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         27-MAR-2000 (first entry)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         Conservative
   Kaur B,
                                                          WPI; 2000-116417/10.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    Local Similarity
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        371 AA;
                                                                                               N-PSDB; AAZ29859
Doetsch PW,
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      Matches 205;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     Sequence
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     Query Match
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        AAY44505;
```

The present sequence is delta228-UV damage endonuclease from D. radiodurans. Delta228-UVDE is a 288 amino acid deletion of the N-terminal of the uvel+ gene product. This is expressed in frame with a GST leader sequence to generate a fusion protein. This provides stable endonuclease fragments for cleaving a double-stranded DNA molecule that has a distorted structure resulting from UV radiation damage, a photoproduct, an abasic site, mismatched nucleotide pairing, aplatinum diaduct, an intercalated molecule or alkylation of a nucleotide. Uvelp can be used in compositions for internal or topical application and as a 292 YGIRFLRLSSEMFP---FASHPVHGYKLAPFASEVLAEAGRVAAELGHRLTTHPGQFTQL 348 115 nsdrpevressvramsaharvmdglglartpwn----llllhggkggrgaelaali--- 166 404 KRNYARLSQSCKNRLVLENDDVGWTVHDLLPVCEELNIPMVLDYHHHNICFDPAHLREGT 463 464 LDISDPKLQE---RIANTWKRKGIKQKMHYSEPCDGAVTPRDRRKHRPRVMTLPPCPPDM 520 new truncated ultraviolet damage endonuclease for treatment of skin 349 GSPRKEVVESAIRDLEYHDEL---LSLLKLPEQQNRDAVMIIH--MGGQFGDKAATLERF 403 39; Gaps DB 21; Length 294; 102; Indels 7.5%; Score 262.5; DB 2.32.2%; Pred. No. 1.1e-14; tive 38; Mismatches 102. AAY85657 standard; Protein; 1341 AA. therapeutic agent for skin cancers. Claim 16; Page 60; 133pp; English. 521 -DLMIEAKDKEQAV--FELMRTFK 541 271 pqieveakgkeeaiaalr1mapfk 294 Avery AM; 99WO-US12910. 98US-0088521 99US-0134752 12-FEB-2001 (first entry) Local Similarity 32.2 nes 85; Conservative Dienococcus radiodurans. Kaur B, WPI; 2000-116417/10. (UYEM-) UNIV EMORY. 294 AA; W09963828-A1. 08-JUN-1999; 08-JUN-1998; 18-MAY-1999; Doetsch PW, 16-DEC-1999 AAY85657; Sequence Query Match cancers œ Matches AAY85657 RESULT ŏ pp δ q ŏ g ò qq οy

```
Human Acinus L protein sequence
                                    Homo sapiens.
                                                                            09-APR-1999;
                                                        19-OCT-2000.
                                                                                                                                                                                                                                                                                                                                   Db
ŏ
                                                                                                                                                                                                                                                                                                              qq
                                                                                                                                                                                                                                                                                                                        õ
                                                                                                                                                                                                                             13;
glutathione-S-transferase signal peptide; uvel+ gene product; UV irradiation; DNA damage; UV radiation damage; photoproduct; abasic site; aplatinum diaduct; mismatched nucleotide pairing; nucleotide alkylation; skin cancer.
```

29; inhibitors, regulators or promoters. The polypeptide induces chromatin aggregation within the cell nucleus without inducing DNA fragmentation. The polypeptide can be used in the investigation and treatment of diseases in which apoptoosis is implicated, including AIDS, Alzheimer's disease, Parkinson's disease, cardiac or brain infarction, herpes or adenovirus infection, diabetes and viral hepatitis. The present sequence Chromatin aggregation induction; human; Acinus; AIDS; herpes; diabetes; Alzheimer's disease; Parkinson's disease; cardiac infarction; brain infarction; adenovirus infection; viral hepatitis. aggregation. The polypeptide is a fragment of the human Acinus protein. Included in the invention is a nucleotide sequence encoding the polypeptide, antisense oligonucleotides, antibodies recognising the polypeptide sequence, and apoptosis regulating agents containing the polypeptide, nucleotide sequence, and sense or antisense oligonucleotides. Also included in the invention is a method for screening compounds for their activity as chromatin aggregation involved in DNA fragmentation used for the regulation of apoptosis and screening of potential apoptosis inhibitors 565 lpshrlkgsa----dssssssssssssssssssssssssspdssgsrshsplrskgrdvagarth 620 89 ATEGLARM-----ERRLQRATKRQKKQLEEDGIPVPSVVSRFPTAPYHHKSTNAEERE- 141 681 atppypglgvcepkertstsssvgarrlsgpesaekhvtgrlgpergspkkceaeeaep 740 142 --AKEP----VLKTHSKD-----VEREAEIGVD------DVVKMEPAATNIIEPED 180 741 paatqpqtsetqtshlpeserihhtveekeevtmdtsenrpendv----peppmpiadqv 796 181 AQDAAERGAARPPAVNSSYLPLPWKGRL-----GYACLNTYLRNAKP----- 222 Gaps 1 MPSRKSKAAALDTPQSESSTFSSTLDSSAPSPARNLRRSGRNILQPSSEKDRD-----H 54 55 EKRSGEELAG------RMMGKDANGHCLR--------EGKEQEEGVKM 88 novel polypeptide which induces chromatin aggregation but is not 621. anprgrpkmgsrstsesrsrsrsrsrssassnsrkslspgvsrdsstsytetkdpssggev This invention relates to a polypeptide which induces chromatin Query Match 4.6%; Score 160.5; DB 21; Length 1341; Best Local Similarity 19.6%; Pred. No. 8.8e-05; Matches 153; Conservative 104; Mismatches 315; Indels 209; Disclosure; Page 60-67; 99pp; Japanese. represents the human Acinus protein. Sahara S, Eguchi Y, Tsujimoto Y; 07-APR-2000; 2000WO-JP02254. 99JP-0103317. (SHIO) SHIONOGI & CO LTD. WPI: 2000-665130/64. Sequence 1341 AA; N-PSDB; AAC61196 WO200061743-A1. qq Q D ò

```
| || || gdtltrsisqqksgvsitiddpvrtaqvpsppr----gkisnivhisnlvrpftlgglk 1029
                                                                                                                                                                                   erewdrdkvregprsrsrsrdrrrkerakskekksekkekaqeeppakllddlfrktkaa 1221
              PCDGAVTPRDRRKHRPRVMTLPPCPPDMDLMIEAKDKEQAVFE--LMRTFKLPGFEKIND 550
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    Primer sets for synthesizing polynucleotides, particularly the 5602 full-length cDNAs defined in the specification, and for the detection and/or diagnosis of the abnormality of the proteins encoded by the
----PIFSSRTCRMASIVDHRHPLQFEDEPEHHLKNKPDKSKEPQDEL---GHKFVQEL 274
                                                     GLANARDIVKMLCWNEKYGIRFLRLSSEMFPFASHPVHGYKLAPFASEVLAEAGRVAAEL 334
                                                                                                             GHRLTTH-----SLLKLP 376
                                                                                                                                                                   377 EQQNRDAVMIIHMGGQFGDKAATLERFKRN----YARLSQSCKNRLVLENDDVGWTVHDL 432
                                                                                                                                                                                                                          433 LPVCEELNIPMVLDYHHHNICFDPAHLREGTLDISDPKLQERIANTWKRKGIKQKMHYSE 492
                                                                                                                                                                                                                                                                                                                                                                                              602 PEGPKEVPEEERAMGGPYNRVYW-PL-----GCEEWLKPKKREVKKGKVPEEVEDEG 652
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        Human; primer; detection; diagnosis; antisense therapy; gene therapy
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  Isogai T, Nishikawa T, Hayashi K, Saito K, Yamamoto J;
, Sugiyama T, Wakamatsu A, Nagai K, Otsuki T;
                                                                                 914 glkicrtvtqvvpaegqengqreeeeeekepeaeppvppqvsvevalpppaehevkkvtl
                                                                                                                                                                                                                                                                                                                                        551 MVPYDRD----DENRPAPPVKAPKKKGGKRKRTTDEEAAEPEEVDTAADDV----KDA
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            Human protein sequence SEQ ID NO:15022.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       AAB94420 standard; Protein; 914 AA.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 99JP-0248036.
99JP-0300253.
2000JP-0118776.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      28-JUL-2000; 2000EP-0116126
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            02-MAY-2000; 2000JP-0183767
09-JUN-2000; 2000JP-0241899
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                (first entry)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       (HELI-) HELIX RES INST.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             WPI; 2001-318749/34.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  29-JUL-1999;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                27-AUG-1999;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    Homo sapiens
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               EP1074617-A2
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              11-JAN-2000;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              26-JUN-2001
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           07-FEB-2001
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  1264 e 1264
                                                                                                                                                                                                                                                                                                                                                                                                                           1222 p----
                                                                                                                                                                                                                                                                                                                                                                                                                                                       653 E 653
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                Ishii S,
                                                                                                                                                                                              1030
                                                                                                                                                                                                                                                                                                           1122
                                                                                                                                                                                                                                                                                                                                                                  1162
                                                     275
                                                                                                            335
                                                                                                                                                                                                                                                                                 493
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         AAB94420
                                                                                                                                                                                                                                                                                                          g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     g
                                                                                                                                     g
                                                                                                                                                                                              Pb
                                                                                                                                                                                                                          ð
                                                                                                                                                                                                                                                   Q
                                                                                                                                                                                                                                                                                                                                      ò
                                                                                                                                                                                                                                                                                                                                                                  g
                                                                                                                                                                                                                                                                                                                                                                                              ò
                                                                                                                                                                                                                                                                                                                                                                                                                        q
                                                                                g
                                                                                                            δ
                                                                                                                                                                   ò
                                                                                                                                                                                                                                                                                 ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                       õ
                                                     ò
```

```
to the complementary strain of a polynucleotide which comprises one of the 5602 nucleotide sequences defined in the specification, where the oligonucleotide comprises at least 15 nucleotides; or (b) a combination of an oligonucleotide comprising a sequence complementary to the complementary strand of a polynucleotide which comprises a 5'-end sequence complementary to the complementary strand of a polynucleotide which comprising a sequence complementary to a polynucleotide which comprises a 3'-end sequence complementary to a polynucleotide which comprises a 3'-end sequence of the combination of the 5'-end sequence/3'-end sequence is selected from those defined in the specification. The primer sets can be used in antisense therapy and in gene therapy. The primers are useful for synthesising polynucleotides, particularly full-length cDNAs. The primers are also useful for the full-length cDNAs. The primers also useful for the full-length cDNAs. The primers also useful for the convasion of the full-length cDNAs as allow obtaining of the full-length cDNAs are primers allow obtaining of the full-length cDNAs are primers allow obtaining of the full-length cDNAs are primers allow as an anomal and AAH13629 to AAH13632 represent oligonucleotides, all of which are used in the exemplification of the present invention.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           39;
                                                                                              The present invention describes primer sets for synthesising 5602 full-length cDNAs defined in the specification. Where a primer set comprises: (a) an oligo-dT primer and an oligonucleotide complementary
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       64 GRMMGK---DANGHCLREGKEQEEGVKMAI-----EGLARMERRLQR--ATKRQKKQLE 112
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             EDGIPVPSVVSRFPTAPYHHKSTNAEEREAKEPVLKTHSKDVEREAEJGVDDVVKMEPAA 172
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        pggttpkgll-----pdsfhtetledtekpsvnerpsekdaeqegegg---giitrgtl 269
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  173 TNIIEPEDAQDAAERGAARPPAVNSSYLPLPWKGRLGYACLNTYLRNAKPPIFSSRTCRM 232
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         233 ASIVDHRHPLQFEDEPEH-------HIKNKPDKSKEPQDELGHKFVQELGL 276
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  321 esiackrdpeqakeepghiptghveeavnevdnentqqkddesdapcsslg-----1ss 374
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   277 ANARDIVKMLCWNEKYGIRFLRLSSEMFPFASHPVHGYKLAPF-ASEVI,AEAGRVAAELG 335
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  375 sgsgdaara---pksvprpkrltsstynlr----hahslgsldaskvtse--keaagvn 424
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       336 HRLTTHPGQFTQLGSPRKE-----VVESAIRDLEYHDELLSLLKLPEQQNRDAVMIIHMG 390
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      pimpkenga-sesgdpldeddvdtvvdegpkfmew-----caeeenge--lianfn 472
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          ------GDKAATLERFKRNYARLSQSCKNRLVLENDDVGW 427
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         473 agymkvgkgwiglekeggptprarnksdklkeiwkskkr---srkcrsslesgk---- 523
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          -----cspvqmlfmtnfklsnvckwflettetrslvivkklntrlpggvppvkhplg 575
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              454 -FDPAHLREGTLDISDPKLQERI-----ANTWKRKGIKQKMHYS----EPCDGA-V 498
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                11 LDTPQSESSTF----SSTLDSSAPSPARNLRRSGRNILQPSSEKDR-DHEKRSGEE-LA 63
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       4.1%; Score 143.5; DB 22; Length 914;
19.3%; Pred. No. 0.0016;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           Indels
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           281;
                                               Claim 8; SEQ ID 15022; 2537pp + CD ROM; English.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           Mismatches
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           428 TVHDLLPVCEELNIPMVLDYHHNIC----
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        Conservative 103;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              19.3%;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              Best Local Similarity
Matches 156; Conserv
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             914 AA;
  full-length cDNAs
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       391 GOF----
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                Sequence
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         Query Match
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              113
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                219
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        425
ő
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         Dp
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             qq
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              Q
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           Op
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              q
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                ŏ
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            Ωp
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   ÓΥ
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                QQ
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   QQ
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      q
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        QΥ
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              Óχ
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             qq
```

```
The present sequence represents a human microtubule-associated protein designated TPX2. TPX2 functions by assisting the binding of kinesin-like proteins having a motor function to microtubules during mitosis. TPX2 is a receptor for a leucine zipper of the kinesin-like protein. The TPX2 nucleic acid can be used directly to induce or promote mitosis, or to compensate for cell defliciency. The TPX2 polypeptide is useful for production of an antibody. The TPX2 polypeptide is useful is useful for inducing or inhibiting mitosis. In particular, the inhibitor is useful for treating cancer.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  Microtubule-associated protein; TPX2; kinesin-like protein; microtubule;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     New nucleic acids encoding microtubule-associated targeting proteins for Xklp2 (TPX2), obtained from Xenopus laevis and humans, useful for modulating mitosis and for producing TPX2 inhibitors for treating
                                 634 spgpnsedsieevkedrnshppanlptpastrilrkysnirgklraggrliknekmecpd 693
                                                                                                          694 --alaveskpsrksvc-----inplmspklalqvdaddfpvkpkstegmkgr 738
                                                                                                                                            569 ------PKKKKGGKRKRTTDEEAAEPEEVDTAADDVKDAPE----GPKEVPEEERAM 615
-----PCPP 518
                                                                                                                                                                                739 kgkgvseilpkaevgskrkrt---egssppdsknkgptvkaskekhadgatktpaakrp- 794
                                                                       519 DMDLMIEAKDKEQAVFELMRTFKLPGFEKINDM----VPYDRDDENRPAPPVKA-----
                                                                                                                                                                                                                                                                                                                                                                                                                                                                   A human microtubule-associated protein designated TPX2.
499 TP--------RDRRKHRPRVMTLP---------
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                (EMBL-) EMBL EURO LAB MOLEKULARBIOLOGIE.
                                                                                                                                                                                                                     616 GGPYNRVYWPLGCEEWLKPKKREVKKGKV 644
                                                                                                                                                                                                                                                           795 -aardrssqp-----pkktslkenkv 814
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  Claim 13; Page 36-39; 54pp; English.
                                                                                                                                                                                                                                                                                                                                                    AAB19002 standard; Protein; 747 AA
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          99EP-0106710.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     99EP-0110507
                                                                                                                                                                                                                                                                                                                                                                                                                               (first entry)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                WPI; 2000-595692/57.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            747 AA;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     N-PSDB; AAA96602
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          mitosis; cancer
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               EP1041147-A2
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       31-MAY-1999;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            01-APR-1999;
                                                                                                                                                                                                                                                                                                                                                                                                                             08-FEB-2001
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     04-OCT-2000.
                                                                                                                                                                                                                                                                                                                                                                                        AAB19002;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              Sequence
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                cancer
                                                                                                                                                                                                                                                                                                                                  AAB19002
                                                                       ΟŊ
                                                                                                          qq
                                                                                                                                                                                g
                                                                                                                                                                                                                     ò
                                                                                                                                                                                                                                                           g
                                                                                                                                                                                                                                                                                                                                                                        ŏ
```

```
283 VKMLCWNEKYGIRFLRLSSEMFPFASHPVHGYKLAPFASEVLAEAGRVAAELGHRLTTHP 342
                                                                                                                                                                                                                                                                                                                                                                                   25 vemrkkneef------242
                                                                                                                                                                                                                                                                                                                                                                                                                            343 GOFTOLGSPRKEVVESAIRDLEYH---DELLSLLKLPEQONRDAVMIIHMGGQFGDKAAT 399
                                                                                                                                                                                                                                                                                                                                                                                                                                                400 LERFKRNYARLSQSC------KNRLVLENDDVGWTVHDLLPVCEELNIPMVLDYH 448
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       287 lrkhpssparvtkgctivkpfnlsggkkrtfde----tvstyvplagg----vedfh 335
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    336 krtp--nryhlrskkddin------11pskssvtkicrdpqtpvlqtkhra 378
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       509 RVMTLPPCPPDMDLMTEAKDKEQAVFELMRTF------KLPGFEKINDMVPYD-555
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           379 ravt---ckstaeleaeeleklqqykfkareldprileggpilpkkppvkpptepigfdl 435
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  556 -----PVKAPKKKKGGKRKR- 579
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          580 --TIDEEAAEPEEVOTAADDVK--DAPEGPKEVPEEERAMGGPYNRVYWPLGCEEWLKPK 635
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 99 acssleveaaisrktpagpgrrs1rlsagkdlegkekhhvkmkakrcatpviideilpsk 158
                                                                                                                                                                                            168 MEPAATNIIEPEDA----ODAAERGAARPPAVNSSYLPLPWKGRLGYACLNTYLRNAKPP 223
                                                                                                                                                                                                                                                                   224 IFSSRTCRMASTVDHRHPLQFEDEPEHHLKNKPDKSKEPQD-ELGHKFVQELGLANARDI 282
                                                                                                                                                                                                                                                                                                          ----akgkflksteegeleksmkmgge-----v 225
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              449 HHNICFDPAHLREGTLDISDPKLQERIANTWKRKGIKQKMHYSEPCDGAVTPRDRRKHRP 508
                                      63 AGRMMGKDANGHCLREGKE--QEEGVKMAI-EGLARMERRLQRATKRQKKQLEEDGIP-- 117
                                                                                                                 118 -----VPSVVSRFPTAPYHHKSTNAEEREAKEPVLKTHSKDVERE--AEIGVDDVV--K 167
                                                              2 sqvkssysydap-sdfinfss-lddegdt------qnidswfeek-----anl 41
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        Human foetal liver; fls353; recombinant production; antibody;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        AAB29444 standard; Protein; 747 AA.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      Human foetal liver fls353 protein.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          drug screening; anticancer agent.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              07-APR-2000; 2000WO-JP02281.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     99JP-0103356.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               09-FEB-2001 (first entry)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        | :|::||:||:
553 kkikelqkgevpk 565
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    636 K -- REVKKGKVPE 646
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     WO200061744-A.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     09-APR-1999;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   sapiens.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            19-OCT-2000
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          AAB29444;
                                                                                                                                                                                                                                                                                                            202
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   Ношо
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    AAB29444
                                                                                                                                                                                                                                                                       ŎΣ
                                                                                                                                                                                                                                                                                                                                                                                     qq
                                                                                                                                                                                                                                                                                                                                                                                                                                                              qq
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     QΫ́
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       q
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              òγ
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    q
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           q
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    ŏλ
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           ŏ
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       QQ
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            0.7
                                                                                                                                                                                                                                                                                                                                                  6
                                                                                                                                                                                                                                                                                                                                                                                                                            Öλ
```

Gaps

97; Mismatches 234; Indels 258;

Query Match 4.0%; Score 139; DB 21; Length 747; Best Local Similarity 19.6%; Pred. No. 0.0029;

Matches 144; Conservative

3 SRKSKAAALDTPQSESSTFSSTLDSSAPSPARNLRRSGRNILQPSSEKDRDHEKRSGEEL 62

<u>..</u> __ ..

δλ

The invention relates to the novel human foetal liver proteins f1s353 (AAB29444) and f1s485 (L variant (AAB29445) and S variant (AAB29446), and to nucleic acids encoding them (AAC63860-C63862). The invention also relates to expression vectors and host cells comprising an f1s353 or f1s485 DNA, the recombinant production of the proteins, antibodies against the proteins, and a method for screening compounds for their proteins. The f1s353 and f1s485 proteins and nucleic acids can be used to proteins. The f1s353 and f1s485 proteins and nucleic acids can be used to identify compounds which regulate the activity and expression of the proteins. Such compounds may be used as anticancer agents. The present 36; Genes expressed specifically in fetal cells and some cancer cells for 167 168 MEPAATNIIEPEDA----QDAAERGAARPPAVNSSYLPLPWKGRLGYACLNTYLRNAKPP 223 201 225 342 242 336 krtp--nryhlrskkddin-------11pskssvtkicrdpqtpvlqtkhra 378 ravt---ckstaeleaeelekiqqykfkareldprileggpilpkkppvkpptepigfdl 435 63 AGRAMGKDANGHCLREGKE--QEEGVKMAI-EGLARMERRLQRATKRQKKQLEEDGIP-- 117 Gaps sqvkssysydap-sdfinfss-lddegdt-----qnidswfeek-----anl -----VPSVVSRFPTAPYHHKSTNAEEREAKEPVLKTHSKDVERE--AEIGVDDVV--K 224 IFSSRTCRMASIVDHRHPLQFEDEPEHHLKNKPDKSKEPQD-ELGHKFVQELGLANARDI 283 VKMLCWNEKYGIRFLRLSSEMFPFASHPVHGYKLAPFASEVLAEAGRVAAELGHRLTTHP ----iggpvkksvsqvtksvdfhfrtderikghpkngeeykev-----nftse 400 LERFKRNYARLSQSC----KNRLVLENDDVGWTVHDLLPVCEELNTPMVLDYH 449 HHNICFDPAHLREGTLDISDPKLQERIANTWKRKGIKQKMHYSEPCDGAVTPRDRRKHRP 3 SRKSKAAALDTPQSESSTFSSTLDSSAPSPARNLRRSGRNILQPSSEKDRDHEKRSGEEL ---akqkflksteeqeleksmkmqqe----v 343 GQFTQLGSPRKEVVESAIRDLEYH---DELLSLLKLPEQQNRDAVMIIHMGGQFGDKAAT lrkhpssparvtkgctivkpfnlsggkkrtfde----tvstyvplagg----vedfh ---KLPGFEKINDMVPYD-Indels 258; 4.0%; Score 139; DB 21; Length 747; 19.6%; Pred. No. 0.0029; vemrkkneef----kklalag-97; Mismatches 234; (CHUG-) CHUGAI RES INST MOLECULAR MEDICINE INC RVMTLPPCPPDMDLMIEAKDKEQAVFELMRTF----screening potential anticancer agents Claim 1; Fig 4-5; 111pp; Japanese. sequence represents fls353 Matches 144; Conservative 2000-665131/64 Best Local Similarity 747 AA; N-PSDB; AAC63860 Sequence Query Match Nezu J, 226 118 202 526 287 g ò g ò ò g ò 유 ŏ a ò g ò Q ŏ q ò g ò ò

This invention describes a novel oligoribo- or oligodeoxyribonucleotide, characterized in that, it hybridizes to mRNA that encodes protein Ki-67 at a physiologically acceptable salt concentration. The oligoribo- or oligodeoxyribonucleotide which is complementary to Ki-67, a protein active at all stages of the cell cycle except G_0, is useful for therapy of illnesses with increased cell proliferation and particularly for treatment of tumnors, autoimmune diseases, scar formation, inflammation, allergy, rheumatic diseases and defence against transplantation. This sequence represents the human cell cycle protein Ki-67 which is described : | | : : | | : | | 2488 kvdmkeeplavsklt-rtsgettqthteptgdsksikafkespkqildpaasvtgsrrql 2546 2428 kekaealedlvgfkelfqtpghteesmtddkitevsckspqpesfktsrsskgrlkiplv 2487 eiekrigereskkktedehfefhsrpcptkiledvvgvpekkvlpitvpkspafalknri 495 580 --TTDEEAAEPEEVDTAADDVK--DAPEGPKEVPEEERAMGGPYNRVYWPLGCEEWLKPK 635 49 EKDRDHEKRSGEELAGRMMGKDANGHCLREG-----KEQEEGVKMAIEGLARME-RRLQR 102 Indels 200; Gaps 5 KSKAAALD-----TP--QSESSTFSSTLDSSAPSPA----RNLRRSGRNILQPSS 48 ycle protein; Ki-67; therapy; cell proliferation; allergy; treatment; autoimmune disease; scar formation; inflammation; Length 3256; Mismatches 296; Score 137; DB 21; Pred. No. 0.038; (BORS-) FORSCHUNGSZENTRUM BORSTEL ZENT MEDIZIN. Deinert Disclosure; Page 19-29; 36pp; German. Ki-67 gene antisense oligonucleotide AAY50976 standard; Protein; 3256 AA. rheumatic disease; transplantation. in the method of the invention. 92; Human cell cycle protein Ki-67. Gerdes J, Boehle A, 3.98; 19.98; 98DE-1022954. 98DE-1022954 (first entry) Matches 146; Conservative 636 K--REVKKGKVPE 646 553 kkikelgkgevpk 565 WPI; 2000-039964/04. 3256 AA; Best Local Similarity N-PSDB; AAZ43872 DE19822954-A1 Homo sapiens 22-MAY-1998; 22-MAY-1998; 10-MAR-2000 25-NOV-1999 cycle AAY50976; Seguence Flad H, Query Match Lumor AAY50976 qq οy á Οy qq δy ------PVKAPKKKKGGKRKR- 579

```
2603 cpktrprkevkeelsaverltgtsggsththkepasgdegikvlkgrakkkpnpveeeps 2662
                                                                                                                                                            2663 rrrprapkekag----pledlagftelsetsghtgesltagkatk---ipcespplevvd 2715
                                                                                                                                                                                                                       2773 tpapaasvtgsrrrprapresaqaiedlagfkdpaaghteesmtddkttkipcksspele 2832
                                                                                                                                                                                                                                                                                                                                                      : | : : | | :: |
2833 dtatsskrrprtraqkvevkeellavgkltqtsgetthtdkepvgegkgtkafkgpakrn 2892
                                                                                                                                                                                                                                                                                                                                                                                                                               2893 vdaedvigsrrqprapkekaqpledl-asfqelsqt--------- 2927
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    ------pghteelangaadsftsapkqtpdsgkplkisrrvlrapk-- 2967
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        187 RGAARPPAVNSSYLPLPWKGRLGYACLNTYLRNAKPPIFSSRTCRMASIVDHRHPLQFED 246
                                                                                                                                                                                                247 ---EPEHHLKNKPDKSKEPQDELGHKFVQELGLANARDIVKMLCWNEKYGIRFLRLSSEM 303
                                                                                                                                                                                                                                                              304 FPFASHPVHGYKLAPFASEVLAE----AGRVAAELGH------RLTTHP----GQFT 346
                                                                                                                                                                                                                                                                                                                                347 QLGSPRKEVVESAIRDLEYHDELLSLLKL-------PEQQN 380
                                                                                                                                                                                                                                                                                                                                                                                                  381 RDAVMIIHMGGO---FGDKAATLERFKRNYARLSQSCKNRLVLENDDVGWTVHDLLPVCE 437
                                                                                                                                                                                                                                                                                                                                                                                                                                                                  -----GTLDISDP-KLQERIANTWKRK 482
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      542 LPGFEKINDMVPYDRDDENRPAPPVKAPKKKGGKRKRTTDEEAAEPEEVDTAADDVKDA 601
103 ATKROKKOLEEDGIPVPSVVSRFPTAPYHHKSTNAEEREAKEPV-----------146
                                                                Peptide #967 encoded by probe for measuring cervical gene expression.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           Probe; human; microarray; gene expression; cervical epithelial cell;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      483 GIKQKMHYSEPCDGAVTPRDRRKHRPRVMT-LPPCPPDMDLMIEAKDKEQAVFELMRTFK
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         AAM14533 standard; Protein; 2665 AA.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                    438 ELNIPMVLDYHHHNICFDPAHLRE---
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             30-JUN-2000; 2000US-0608408.
03-AUG-2000; 2000US-0632366.
21-SEP-2000; 2000US-0234687.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               2000US-0236359.
2000GB-0024263.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             30-JAN-2001; 2001WO-US00670
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               2000US-0180312.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                2000US-0207456.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         12-OCT-2001 (first entry)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     3074 keehklqdsvpenk 3087
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     602 PEGPK---EVPEEE 612
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              cervical cancer.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             WO200157278-A2.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              Homo sapiens.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               04-FEB-2000;
26-MAY-2000;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               27-SEP-2000;
04-OCT-2000;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              09-AUG-2001.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          AAM14533;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       RESULT 13
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          AAM14533
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     qq
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          g
                                                                                                                                                                                                Qγ
                                                                                                                                                                                                                              Q
                                                                                                                                                                                                                                                              δλ
                                                                                                                                                                                                                                                                                              pp
                                                                                                                                                                                                                                                                                                                                  Qγ
                                                                                                                                                                                                                                                                                                                                                               g
                                                                                                                                                                                                                                                                                                                                                                                                  οy
                                                                                                                                                                                                                                                                                                                                                                                                                                  g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                  δ
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    В
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      ö
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     ò
```

```
The present invention relates to human single exon nucleic acid probes (SENP: see AAIIO08-AAIIS4459). The present sequence is a peptide encoded by one such probe. The SENPs are derived from human HeLa cells. The SENPs can be used to produce a single exon microarray, which can be used for measuring human gene expression in a sample derived from human cervical epithelial cells. By measuring gene expression, the probes are therefore useful in grading and/or staging of diseases of the cervix, notably
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   Note: The sequence data for this patent did not form part of the printed specification, but was obtained in electronic format directly from WIPO
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       1072 papapveqleqvdlppgadpdkeaammpagveegssgdqppyldakpptpgasfsqaesn 1131
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             1225 rkseridre-----klkrsnsprgea-----qkllelkmeaekitrtas 1263
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               1308 egprvtrkrlerelgeaaavpttprrg----rppktrrr-adeeeeneakepaet1kppe 1362
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            1174 -----aeaapesqppasedlevdppvaakdkkpnkskrsktpvqaaavsivekpvt 1224
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            899 keeegergelfasrflhssifegdskrlghlerkeedsdfisgriygkgtsegansttds 958
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          93 LAR----MERRLQRATKRQKKQLEEDGIPVPSVVSRFPTAPYHHKSTNA----EEREAKE 144
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           ----VLKTHSKDVEREAEIGVDDVVKMEPAATNIIEPEDAQDAAE 186
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             187 RGAARPPAVNSSYLP-----LPWKGRLGYACLNTYLRNAKPPIFSSRTCRMASI 235
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              236 VDHR------HPLQFEDEPEHHLKNKPDKSKEPQDELGHKFVQELGLANARDIVKMLC 287
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            288 WNEKYGIRFLRLSSEMFPFASH------324
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              325 AEAGRVAAELGHRLTTHPGQFTQLGSPRKEVVESAIRDLEYHDELLSLLKLPEQQNRDAV 384
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            385 MIIHMGGQFGDKAATLERFKRN-----YARLSQSCKNRLVLENDDVGWTVHDLLPVCEEL 439
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              440 NIPMV----LDYHHHNICFDPAHLREGTLDISDPKLQERIANTWKRKGIKQKMHYSEPCD 495
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               496 GAVTPRDRR-----KHRPRV-MTLPPCPPDMDLMIEAKD-----KEQAVF 534
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              296; Indels 242; Gaps
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           52 RDHEKRSGRELAGRMM----GKDAN--GHCLRE-----GKEQEEGVKMAIEG 92
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          3 SRKSKAAALDIPQSESSTFSSTLD----SSAPSP-----ARNLRRSGRNILQPSSEKD 51
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                1132 vdpepdstqplskpagkseeanepkaekpdatadaepdangk---------
                                                                                                                                                             Human genome-derived single exon nucleic acid probes useful for
                                                                                                                                                                                     analyzing gene expression in human cervical epithelial cells -
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            3.9%; Score 136; DB 22; Length 2665; 18.8%; Pred. No. 0.034;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     at ftp.wipo.int/pub/published_pct_sequences.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              95; Mismatches
                                                                                                                                                                                                                                       Claim 27; SEQ ID No 19359; 487pp; English.
                                                                 Hanzel DK, Chen W, Rank DR;
               (MOLE-) MOLECULAR DYNAMICS INC
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           Matches 147; Conservative
                                                                                                                 WPI; 2001-488901/53.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         Best Local Similarity
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              145 P-----
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          2665 AA;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                    cervical cancer.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             Sequence
                                                                 Penn SG,
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   Query Match
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             QQ
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  qq
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 qq
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    δλ
ολ
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            q
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               QΫ
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  δ
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   δλ
```

```
WO200157270-A2
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       09-OCT-2001
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      Homo sapiens
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         04-FEB-2000;
26-MAY-2000;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   30-JUN-2000;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          03-AUG-2000;
21-SEP-2000;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              09-AUG-2001
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   AAM02259;
                                                                                                                      145 p--
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              496
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   1174
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     RESULT 1
AAM02259
                                                               QQ
                                                                                                                      δ
                                                                                                                                                                             Op
                                                                                                                                                                                                                                 δλ
                                                                                                                                                                                                                                                                                     qq
                                                                                                                                                                                                                                                                                                                                               ōλ
                                                                                                                                                                                                                                                                                                                                                                                                   QQ
                                                                                                                                                                                                                                                                                                                                                                                                                                                       οy
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             qq
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    δy
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           qq
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               ōλ
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         q
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                δ
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               qq
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       QQ
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        Öλ
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      δ
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 δŏ
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      The present invention relates to single exon nucleic acid probes (SENP: see AAI31315-AAI57546). The present sequence is a peptide encoded by one such probe. The probes are useful for producing a microarray for predicting, measuring and displaying gene expression in samples derived from human placenta. The probes are useful for antenatal diagnosis of human genetic disorders.
1363 gwrsprsqktaagggppggkkgknepkvdatrpeattevgpqigvkessmepkaaeeeags 1422
                                                                                                                                                                                                    Peptide #987 encoded by probe for measuring placental gene expression.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 :: |: | | | | : :|: | | |: |: | |: |: ||: ||: ||: ||: ||: ||: ||: ||: ||: ||: ||: ||: ||: ||: ||: ||: ||: ||: ||: ||: ||: ||: ||: ||: ||: ||: ||: ||: ||: ||: ||: ||: ||: ||: ||: ||: ||: ||: ||: ||: ||: ||: ||: ||: ||: ||: ||: ||: ||: ||: ||: ||: ||: ||: ||: ||: ||: ||: ||: ||: ||: ||: ||: ||: ||: ||: ||: ||: ||: ||: ||: ||: ||: ||: ||: ||: ||: ||: ||: ||: ||: ||: ||: ||: ||: ||: ||: ||: ||: ||: ||: ||: ||: ||: ||: ||: ||: ||: ||: ||: ||: ||: ||: ||: ||: ||: ||: ||: ||: ||: ||: ||: ||: ||: ||: ||: ||: ||: ||: ||: ||: ||: ||: ||: ||: ||: ||: ||: ||: ||: ||: ||: ||: ||: ||: ||: ||: ||: ||: ||: ||: ||: ||: ||: ||: ||: ||: ||: ||: ||: ||: ||: ||: ||: ||: ||: ||: ||: ||: ||: ||: ||: ||: ||: ||: ||: ||: ||: ||: ||: ||: ||: ||: ||: ||: ||: ||: ||: ||: ||: ||: ||: ||: ||: ||: ||: ||: ||: ||: ||: ||: ||: ||: ||: ||: ||: ||: ||: ||: ||: ||: ||: ||: ||: ||: ||: ||: ||: ||: ||: ||: ||: ||: ||: ||: ||: ||: ||: ||: ||: ||: ||: ||: ||: ||: ||: ||: ||: ||: ||: ||: ||: ||: ||: ||: ||: ||: ||: ||: ||: ||: ||: ||: ||: ||: ||: ||: ||: ||: ||: ||: ||: ||: ||: ||: ||: ||: ||: ||: ||: ||: ||: ||: ||: ||: ||: ||: ||: ||: ||: ||: ||: ||: ||: ||: ||: ||: ||: ||: ||: ||: ||: ||: ||: ||: ||: ||: ||: ||: ||: ||: ||: ||: ||: ||: ||: ||: ||: ||: ||: ||: ||: ||: ||: ||: ||: ||: ||: ||: ||: ||: ||: ||: ||: ||: ||: ||: ||: ||: ||: ||: ||: ||: ||: ||: ||: ||: ||: ||: ||: ||: ||: ||: ||: ||: ||: ||: ||: ||: ||: ||: ||: ||: ||: ||: ||: ||: ||: ||: ||: ||: ||: ||: ||: ||: ||: ||: ||: ||: ||: ||: ||: ||: ||: ||: ||: ||: ||: ||: ||: ||: ||: ||: ||: ||: ||: ||: ||: ||: ||: ||: ||: ||: ||: ||: ||: ||: ||: ||: ||: ||: ||: ||: ||: ||: ||: ||: ||: ||: ||: ||: ||: ||: ||: ||: ||: ||: ||: ||: ||: ||: ||: ||: ||: ||: ||: ||: ||: ||: ||: ||: ||: ||: ||: ||: ||: ||: ||: ||: ||: ||: ||: ||: ||: ||: ||: ||: ||: ||: ||: ||: ||: ||: ||: ||: ||: ||: ||: ||: ||: ||: ||: ||: ||: ||: ||: ||: ||: ||: ||: ||: ||: ||: ||: ||: ||: ||: ||: ||: ||: ||: ||: ||: ||: ||: ||: ||: ||: ||: ||: ||: ||: ||: ||: ||: ||: ||: ||: ||: ||: ||: ||: ||: ||: ||: ||: ||: ||: ||: ||: ||: ||: 
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       840 sreenwsfldw-dsrfanfrnnkdkekvdsaprpipswymkkkkirtdsegkmddkkedh 898
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              52 RDHEKRSGEELAGRMM-----GKDAN--GHCLRE-------GKEQEEGVKMAIEG 92
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   3 SRKSKAAALDTPQSESSTFSSTLD----SSAPSP-----ARNLRRSGRNILQPSSEKD 51
                                                                                         575 -----GKRK----RTTDEEAAEPEEVDTAADDVKDAPEGPKEVPE-EERAMGGP
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     3.9%; Score 136; DB 22; Length 2665;
18.8%; Pred. No. 0.034;
tive 95; Mismatches 296; Indels 242;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       Human genome-derived single exon nucleic acid probes useful for
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         Probe; microarray; human; placenta; antenatal diagnosis;
                                                         535 ELMRTFKLPGFEKINDMVPYDRDDENRPAPPVKAPKKKG-
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    analyzing gene expression in human placenta
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  Claim 27; SEQ ID No 27219; 654pp; English.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               Chen W, Rank DR;
                                                                                                                                                                                                                                                                                                                                                             AAM26950 standard; Protein; 2665 AA
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  (MOLE-) MOLECULAR DYNAMICS INC
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         26-MAY-2000; 2000US-0207456.
30-JUN-2000; 2000US-0608408.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             03-AUG-2000; 2000US-0632366.
21-SEP-2000; 2000US-0234687.
27-SEP-2000; 2000US-0236359.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        30-JAN-2001; 2001WO-US00663
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               04-FEB-2000; 2000US-0180312
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 04-OCT-2000; 2000GB-0024263
                                                                                                                                                                                                                                                                                                                                                                                                                                                                           (first entry)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  Best Local Similarity 18.89
Matches 147; Conservative
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         Hanzel DK,
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                WPI; 2001-488897/53
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       2665 AA;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        genetic disorder
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             WO200157272-A2.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                           17-0CT-2001
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    09-AUG-2001
                                                                                                                                                                                                                                                                                                                                                                                                                       AAM26950;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         Penn SG,
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       Sequence
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        Query Match
                                                         ò
                                                                                                             g
                                                                                                                                                                    ò
                                                                                                                                                                                                                     a
                                                                                                                                                                                                                                                                                                                                                                                                                    qq
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           ŏ
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               g
```

```
Probe; human; breast disease; breast cancer; development disorder; inflammatory disease; proliferative breast disease; non-carcinoma tumour.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        Peptide #941 encoded by probe for measuring human breast gene expression.
                            1072 papapveqleqvdlppgadpdkeaammpagveegssgdqppyldakpptpgasfsqaesn 1131
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         1363 gwrsprsqktaagggpggkkgknepkvdatrpeattevgpqigvkessmepkaaeeeags 1422
                                                                                                                                            1017 ppsvgppsvtvvtlesapsalekttgdktveapl-vteektvepatvs----eeakpase 1071
                                                                                                                                                                                                                                                                                                                                                                                                                                          --aeaapesgppasedlevdppvaakdkkpnkskrsktpvqaaavsivekpvt 1224
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        1225 rkseridre-----klkrsnsprgea------qkllelkmeaekitrtas 1263
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         1264 knsaadlehpepsipisrtrrrnvrsvyatmgdh-enr------spvkepv 1307
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     1423 eqkrdrkdagtdknppetapvevvekkpape-knskskrgrsrnsrlavdksaslknvda 1481
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      1482 avsprgaaagageresgvvavspeksespgkedglssglksdpvdpdkepekedvsasgp 1541
93 LAR----MERRLQRATKRQKKQLEEDGIPVPSVVSRFPTAPYHHKSTNA----EEREAKE 144
                                                                                                 --VLKTHSKDVEREAEIGVDDVVKMEPAATNIIEPEDAQDAAE 186
                                                                                                                                                                                            187 RGAARPPAVNSSYLP-----LPWKGRLGYACLNTYLRNAKPPIFSSRTCRMASI 235
                                                                                                                                                                                                                                                                                                                                                                                         288 WNEKYGIRFLRLSSEMFPFASH------324
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          325 AEAGRVAAELGHRLTTHPGQFTQLGSPRKEVVESAIRDLEYHDELLSLLKLPEQQNRDAV 384
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         385 MIIHMGGQFGDKAATLERFKRN----YARLSQSCKNRLVLENDDVGWTVHDLLPVCEEL 439
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       440 NIPMV----LDYHHHNICFDPAHLREGTLDISDPKLQERIANTWKRKGIKQKMHYSEPCD 495
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        GAVTPRDRR------KHRPRV-MTLPPCPPDMDLMIEAKD-----KEQAVF 534
                                                                                                                                                                                                                                                                                           236 VDHR------HPLQFEDEPEHHLKNKPDKSKEPQDELGHKFVQELGLANARDIVKMLC
                                                                                                                                                                                                                                                                                                                                           1132 vdpepdstqplskpaqkseeanepkaekpdatadaepdangk-------
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       ELMRTFKLPGFEKINDMVPYDRDDENRPAPPVKAPKKKG-----------
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     -----GKRK----RTTDEEAAEPEEVDTAADDVKDAPEGPKEVPE-EERAMGGP
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           AAM02259 standard; Protein; 2665 AA
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           2000US-0608408.
2000US-0632366.
2000US-0234687.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               2000US-0180312.
2000US-0207456.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  2000US-0236359.
2000GB-0024263.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            29-JAN-2001; 2001WO-US00661.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               (first entry)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  27-SEP-2000;
04-OCT-2000;
```

us-09-724-296-36.rag

Ouery Match 3.9%; Score 136; DB 22; Length 2665; Best Local Similarity 18.8%; Pred. No. 0.034; Matches 147; Conservative 95; Mismatches 296; Indels 242;

31; 1225 rkseridre-----klkrsnsprgea-----qkllelkmeaekitrtas 1263 ----spvkepv 1307 1072 papapveqleqvdlppgadpdkeaammpagveegssgddppyldakpptpgasfsqaesn 1131 1174 -----aeaapesgppasedlevdppvaakdkkpnkskrsktpvgaaavsivekpvt 1224 385 MIIHMGGQFGDKAATLERFKRN----YARLSQSCKNRLVLENDDVGWTVHDLLPVCEEL 439 440 NIPMV----LDYHHHNICFDPAHLREGTLDISDPKLQERIANTWKRKGIKQKMHYSEPCD 495 145 P------VLKTHSKDVEREAEIGVDDVVKMEPAATNIIEPEDAQDAAE 186 187 RGAARPPAVNSSYLP-----LPWKGRLGYACLNTYLRNAKPPIFSSRTCRMASI 235 236 VDHR-----HPLQFEDEPEHHLKNKPDKSKEPQDELGHKFVQELGLANARDIVKMLC 287 ----- 324 325 AEAGRVAAELGHRLTTHPGQFTQLGSPRKEVVESAIRDLEYHDELLSLLKLPEQQNRDAV 384 899 keeegergelfasrflhssifegdskrlghlerkeedsdfisgriygkgtsegansttds 958 93 LAR----MERRLORATKROKKOLEEDGIPVPSVVSRFPTAPYHHKSTNA----EEREAKE 144 840 sreenwsfldw-dsrfanfrnnkdkekvdsaprpipswymkkkkirtdsegkmddkkedh 898 52 RDHEKRSGEELAGRMM-----GKDAN--GHCLRE------GKEQEEGVKMAIEG 92 3 SRKSKAAALDTPQSESSTFSSTLD----SSAPSP-----ARNLRRSGRNILQPSSEKD 51 1132 vdpepdstqplskpaqkseeanepkaekpdatadaepdangk-----1264 knsaadlehpepslplsrtrrrnvrsvyatmgdh-enr---288 WNEKYGIRFLRLSSEMFPFASH-----qq q QQ βD δ δλ Ω Op δy q δ ò

534	1422	574	1481	618	1541
496 GAVTPRDRRKRRPRV-MTLPPCPPDMDLMIEAKDKEQAVF 534	аадддрадкка			GKRKRTTDEEAAEPEEVDTAADDVKDAPEGPKEVPE-EERAMGGP 618	1482 avsprgaaaqageresgvvavspeksespgkedglssgiksdpvdpdkepekedvsasgp 1541
496	1363	535	1423	575	1482
Οy	qq	Οy	qq	QY	qq

Search completed: January 15, 2002, 13:50:07 Job time: 296 sec

```
GenCore version 4.5 Copyright (c) 1993 - 2000 Compugen Ltd.
```

OM protein - protein search, using sw model

January 15, 2002, 13:52:23; Search time 74.94 Seconds (without alignments) 666.807 Million cell updates/sec Run on:

US-09-724-296-36 Title: Perfect score: Sequence:

1 MPSRKSKAAAALDIPQSESST.....REVKKGKVPEEVEDEGEFDG 656

BLOSUM62 Gapop 10.0 , Gapext 0.5 Scoring table:

219241 seqs, 76174552 residues Searched:

219241 Total number of hits satisfying chosen parameters:

Minimum DB seq length: 0 Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0% Maximum Match 100% Listing first 45 summaries

Database :

PIR_68:*
1: pirl:*
2: pir2:*
3: pir3:*
4: pir4:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

Description	Uv-endonuclease - UV-endonuclease - UV-endonuclease - Nypothetical prote microtubule-associ hypothetical prote nypothetical prote NF-180 - sea lampt hypothetical prote nypothetical prote NF-180 - sea lampt hypothetical prote lyA-specific metal blackjack protein, proliferation neurofilament trip Bassoon protein - hypothetical prote call proliferation neurofilament trip Bassoon protein - hypothetical prote caldesmon - human centromere protein in hypothetical prote CESI protein - hypothetical prote CESI protein - hypothetical prote hypothetical prote hypothetical prote hypothetical prote hypothetical prote	
SUMMARIES	S55262 S71134 E83974 S55418 C555418 C75350 T00365 T17292 T17292 T33100 C81169 T42727 T42730 T42730 T42730 T42730 T42730 T42730 T22466 S28261 ES5485 T22490 A43427	
DB		
ĝ	556 556 556 326 326 326 326 1280 536 1110 1110 1110 1110 1110 1110 1110 11	
ery tch	1 0 0 0 0 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1	
cor	1011.5 377 348.5 361.5 160.5 160.5 1140 141.5 139.5 139.5 137 137 137 136.5 136.5 136.5 136.5 137 137 137 136.5 137 137 136.5 137 136.5 137 137 137 137 137 136.5 137 137 137 137 137 137 137 137 137 137	
n Z	20 20 20 20 20 20 20 20 20 20 20 20 20 2	

elastic titin - hu	neurofilament trin	The rest of the control of the contr	hypothotical	arpoint for profe	groovin gene prote	hucethotical receptor c	adonomateur prote	adenomacous polypo	Tay coositi	hinothetical	hypothetical prote	hypumetical prote	busetbetien	nypormerical proce	ankyiin 3, long sp neurofilament trip
138346	QFHUH	A34736	T34418	T13734	860254	T26819	713825	C96608	A81937	C86406	T52359	071623	726135	A55575	A37221
7	П	~	7	7	2	O	7	7	7	7	~	2	N	100	-
7962	1020	1805	3488	4151	2453	720	2416	522	1773	786	1400	2441	1829	4377	1072
3.7	3.7	3.7	3.7	3.7	3.7	3.6	3.6	3.6	3.6	3.6	3.6	3.6	3.6	3.6	3.6
128.5	128	128	128	128	127.5	126.5	126.5	126	126	125.5	125.5	125.5	124.5	124.5	124
30	3.1	32	33	34	35	36	37	38	39	40	41	42	43	44	45

ALIGNMENTS

RESULT S55262

421 ENDDVGWTVHDLLPVCEELNIPMVLDYHHHNICFDPAHLREGTLDISDPKLQERIANTWK 480

δλ

```
G.; Sasaki, R.; Masui, N.; Fuji, F.; H
                                                                                                                                                                                                                                                                                                                                                                    phillic bacterium Bacillus haloduruns a
                                                                                                                                                                                                                                                                                                                                                                                                                                               NID:910175192; PIDN:BAB06316.1; GSPDB:G
                                                                                                                                                                                                                                                                                                                    0000 #text_change 31-Dec-2000
                                                                                                                                                                                                                                                                                             us halodurans (strain C-125)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         DB 2; Length 322;
4e-16;
s 127; Indels 68; Gaps
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               IDLEYHDELLSLLKLPEQONRDAVM 385
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 NDDVGWTVHDLLPVCEELNIPMVL 445
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               KGIKQKMHYSEPCDGAVTPR---- 501
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    QAVFELMRTFKLPGFEKINDMVPY 554
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         |:|:||:
|LALFQLMK------DLAEY 307
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             QFEDEPEHHLKNKPDKSKEPQDEL 266
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 MFPFASHPV-HGYKLAPFASEVLA 325
                                                QFEKIEDH-----EA 36
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              : | :||: |:|
LVPLVNHPLTEGWKYELAIAEELQ 96
                                                                                                                                                     KKKKGGKRKRTTDEEAAEPEEVD 592
SQFGDKAATLERFKRNYARLSQS 413
              GTFEGKKETLDRFRKNYQRLSDS 426
                                                                                                                                                                                                      SEWLKPKKREVKKGKVPEEVEDE 651
                                                                                                                                                                                                                         : | :|| :| | ||||::
SKRLTARKRRSRK----EEVEED 597
```

2

13;

```
A:Cross-references: GB:AE002022; GB:AE000513; NID:g6459590; PIDN:AAF11370.1; PID:g645
A:Experimental source: strain R1
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   R;Ishikawa, K.; Nagase, T.; Suyama, M.; Miyajima, N.; Tanaka, A.; Kotani, H.; Nomura, DNA Res. 5, 169-176, 1998
A;Title: Prediction of the coding sequences of unidentified human genes. X. The compl A;Reference number: 214142; MUID:98403880
                            R;White, O.; Eisen, J.A.; Heidelberg, J.F.; Hickey, E.K.; Peterson, J.D.; Dodson, M.; Shen, M.; Vamathevan, J.J.; Lam, P.; McDonald, L.; Utterback, T.; Zalewski, S.; Smith, H.O.; Venter, J.C.; Fraser, C.M. Science 286, 1571-1577, 1999
A;Title: Genome sequence of the radioresistant bacterium Deinococcus radiodurans RA;Reference number: A75250; MUID:20036896
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           A;Residues: 1-1280 <15H>
A;Cross-references: EMBL:AB014570; NID:93327153; PIDN:BAA31645.1; PID:93327154
A;Experimental source: brain; clone HK02359
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       C;Species: Homo sapiens (man)
C;Date: 01-Feb:1999 #sequence_revision 01-Feb-1999 #text_change 21-Jul-2000
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      292 YGIRFLRLSSEMFP---FASHPVHGYKLAPFASEVLAEAGRVAAELGHRLTTHPGQFTQL 348
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             145 NSDRPEVRESSVRAMSAHARVMDGLGLARTPWN-----LLLLHGGKGGRGAELAALI--- 196
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 197 ----PDLPDPVRLRLGLENDERAYSPAELLPICEATGTPLVFDAHHHVV------HDKL 245
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   246 PDQEDPSVREWVLRARATWQPPE-WQVVHLSNGIEG---PQDRR-HSHLIADFPSAYADV 300
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         86 HDIRLYRLSSSLFPMLDLAGDDTGAAVLTHLAPQLL.-EAGHAFTDAGVRLLMHPEQFIVL 144
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             349 GSPRKEVVESAIRDLEYHDEL---LSLLKLPEQQNRDAVMIH--MGGQFGDKAATLERF 403
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 404 KRNYARLSQSCKNRLVLENDDVGWTVHDLLPVCEELNIPMVLDYHHHNICFDPAHLREGT 463
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            464 LDISDPKLQE----RIANTWKRKGIKQKMHYSEPCDGAVTPRDRRKHRPRVMTLPPCPPDM 520
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        Indels 209; Gaps
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             1 MPSRKSKAAALDTPQSESSTFSSTLDSSAPSPARNLRRSGRNILQPSSEKDRD-----H 54
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        ----EGKEQEEGVKM 88
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             Length 1280;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             Length 326;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             DB 2;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      Query Match

4.6%; Score 160.5; DB 2;
Best Local Similarity 19.6%; Pred. No. 0.042;
Matches 153; Conservative 104; Mismatches 315;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         Mismatches 102;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  7.5%; Score 261.5; DB 2 32.2%; Pred. No. 3.7e-09;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 55 EKRSGEELAG------RMMGKDANGHCLR-----
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              A;Status: preliminary; translated from GB/EMBL/DDBJ
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   hypothetical protein KIAA0670 - human (fragment)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            521 D-LMIEAKDKEQAV--FELMRTFK 541
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     301 PWIEVEAKGKEEAIAALRUAPFK 324
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      38;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         Conservative
                                                                                                                                                                                                                                                                                                                                           A; Residues: 1-326 <WH1>
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      Best Local Similarity
                                                                                                                                                                                                                                                                       A;Status: preliminary A;Molecule type: DNA
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  A; Molecule type: mRNA
C; Accession: C75350
                                                                                                                                                                                                                                       A; Accession: C75350
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  T00365
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          T00365
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              A; Map position: 1
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      85;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      A; Note: KIAA0670
                                                                                                                                                                                                                                                                                                                                                                                                                                                                              A; Gene: DR1819
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      Query Match
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       C; Accession:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  A; Accession:
                                                                                                                                                                                                                                                                                                                                                                                                                                              C; Genetics:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      Matches
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             δλ
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         QQ
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             ŏ
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             рp
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 δ
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    qq
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        δŏ
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        qq
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        Qγ
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            QO
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             qq
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 ŏ
                                                                                                                                                                                                                                                                                                                                                                                                                                              A.Rolecule type: DNA
A.Rolecule type: DNA
A.Rolecule type: DNA
A.Rolecule type: DNA
A.Rolecule type: DNA
A.Rolecule type: DNA
A.Rolecule type: DNA
A.Rolecule type: DNA
A.Rolecule type: DNA
A.Cooss-references: EMBL: 249782; NID: 9853752; PIDN: CAA89865.1; PID: 9853757
B.Kunst. F.; Ogasawara, N.; Moszer, I.; Albertini, A.M.; Alloni, G.; Azevedo, V.; Berter
C.; Bron, S.; Brouillet, S.; Bruschi, C.V.; Cadduni, B.; Capuano, V.; Carter, N.M.; Chc
A.; Ehrlich, S.D.; Emmerson, P.T.; Entian, K.D.; Errington, J.; Fabret, C.; Ferrari, E.
Nature 390, 249-256, 1997
A.Authors: Foulger, D.; Fritz, C.; Fujita, M.; Fujita, Y.; Fuma, S.; Galizzi, A.; Galler
iech, J.; Harwood, C.R.; Henaut, A.; Hilbert, H.; Holsappel, S.; Hosono, S.; Hullo, M.F.
Koetter, P.; Koningstein, G.; Krogh, S.; Kundin, M.; Murite, R.; Laddus, A.; Ladarevic, V.; Each, S.M.; Levine, A.; Liu, H.; Masuda, S.; Maueel
Y. M.; Ogawa, K.; Ogiwara, A.; Oudega, B.; Park, S.H.; Parro, V.; Pohl, T.M.; Portetelle
Rieger, M.; Rivolta, C.; Rocha, E.; Roche, B.; Rose, M.; Sadale, Y.; Sato, T.; Sato,
A.Authors: Schleich, S.; Schroeter, R.; Scoffone, F.; Sekiguchi, J.; Sekowska, A.; Sekor,
A.; Winters, P.; Winters, P.; Wamane, K.; Yasumoto, K.; Yata, K.; Yoshida, R.; Altile: The complete genome sequence of the Gram-positive bacterium Bacillus subtilis.
A.Reference number: Apples
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        A;Cross-references: GB:Z99123; GB:AL009126; NID:g2636240; PIDN:CAB15748.1; PID:e1186221; A;Experimental source: strain 168 C;Genetics:
                                                                                                                                                                                                                                                     RiGlaser, P.; Danchin, A. submitted to the EMBL Data Library, May 1995
A; Description: Cloning and sequencing of the Bacillus subtilis chromosomal region from
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                8
                                                                                                                                                     C;Species: Bacillus subtilis
C;Date: 15-Jul-1995 #sequence_revision 01-Sep-1995 #text_change 15-Oct-1999
C;Accession: S55418; B70060
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           probable UV damage endonuclease - Deinococcus radiodurans (strain R1)
C;Species: Deinococcus radiodurans
C;Date: 03-Dec-1999 #sequence_revision 03-Dec-1999 #text_change 31-Mar-2000
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              105 LRISFHPNQFTLFISPKESVIKNAVIDMAXHYRMLEAMGIADR----SVINIHIGGAYGN 160
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         277 ANARDIVKMLCWNEKYGIRFLRLSSEMFPFASHP-VHGYKLAPFASEVLAEAGRVAAELG 335
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            336 HRLTTHPGQFTQLGSPRKEVVESAIRDLEYHDELLSLLKLPEQQNRDAVMIIHMGGQFGD 395
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           396 KAATLERFKRNYARLSQSCKNRLVLENDDVGWTVHDLLPVCEELNIPMVLDYHHHNICFD 455
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               Gaps
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    456 PAHLREGTLDISDPKLQERIANTWKRKGIKQKMHYSEP-CDGAVTPRDRRKHR----
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            Length 320;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          Status: nucleic acid sequence not shown; translation not shown
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     10.0%; Score 348.5; DB 2; Length 32.0%; Pred. No. 1.4e-14; tive 52; Mismatches 104; Indels
                                                                                                                         UV-endonuclease homolog ywjD - Bacillus subtilis
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      1 : ::| ||||| |::|: ||
268 LLPLLERFRQWGTNIDFMIEAKQKDKALLRLM 299
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        508 -- PRVMTLPPCPPDMDLMIEAKDKEQAVFELM 537
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     Local Similarity 32.0%
les 87; Conservative
                                                                                                                                                                                                                                                                                                                                                                    A; Reference number: S55414
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             A; Residues: 1-320 <KUN>
                                                                                                                                                                                                                                                                                                                                                                                                                                    A; Status: preliminary
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             A; Molecule type: DNA
                                                                                                                                                                                                                                                                                                                                                                                                A; Accession: S55418
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          A; Accession: B70060
A; Status: nucleic ac
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     Query Match
Best Local S
Matches 87
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 A;Gene: ywjD
```

29;

C75350

q

δ

q

õ

ò

ò

qq	260	ANPRGRPKMGSRSTSESRSRSRSRSRSSSNSRKSLSPGVSRDSSTSYTETKDPSSGQEV 619	6
Οy	83	AIEGLARMERRLQRATKROKKQLEEDGIPUPSVVSRFPTAPYHHKSTNAEERE- 141	1
qq	620	ATPPVPQLQVCEPKERTSTSSSSVQARRLSQPESAEKHVTQRLQPERGSPKKCEAEEAEP 679	6
Qy	142	AKEPVLKTHSKDVEREAEIGVDDVVKMEPAATNIIEPED 180	0
qq	089	PAATQPQTSETQTSHLPESERIHHTVEEKEEVTMDTSENRPENDVPEPPMPIADQV 735	r.
Qy	181	22	2
qq	736	SNDDRPEGSVEDEEKKESSLPKSFKRKISVVSATKGVPAGNSDTEGGQPGRKRRWGASTA 795	τ.
οy	223	DEPEHHLKNKPDKSKEPQDELGHKFVQEL : : : :	4
Dp	796		2
yo 4	275	GLANARDIVKMLCMNEKYGIRFLRLSSEMFPFASHPVHGYKLAPFASEVLAEAGRVAAEL 334 [4 0
3 3	, ,		ט ו
Dp Dp	913	SISQQKSGVSITIDDPVRTAQVPSPPRGKISNIVHISNIVRPFTLGQLK	o &
Qy	377	DVGWTVHDL 43	2
QQ	696	ELLGRIGILVEEAFWIDKIKSHCFVIYSTVEEAVATRIALHGVKWPQSNP 101	18
٥y	433	LPVCEELNIPMVLDYHHHNICFDPAHLREGTLDISDPKLQERIANTWKRKGIKQKMHYSE 492 : : : : : : :	7
QQ	1019	KFLCADYAEQDELDYHRGLLVDRPSETKTEEQGIPRPLHPPP 106	09
Qy	493	PCDGAVTPRDRRKHRPRVMTLPPCPPDMDLMIEAKDKEGAVFELMRTFKLPGFEKIND 550	00
Qy	551	MVPYDRDDENRPAPPVKAPKKKGGKRKRTTDEEAAEPEEVDTAADDVKDA 601	
Dp	1101	EREMDROKVREGPRSRSRSRDRRRKERAKSKEKKEKAQEEPPAKLLDDLFRKTKAA 1160	09
Qy	602	-PLGCEEWLKPKKREVKKGKVPEEVEDEG 65	2
Db	1161	PCIYWLPLTDSQIVQKEAERAKEREKRKKEQEEEEQKER 120	02
Qy	653	E 653	
Dp	1203	F 1203	
RESULT 13564 N.) Spice N.) Spice C.) Spice C.) Date C.) Span Submit A.) Stath A.) Stath A.) Stath A.) Stath A.) Stath A.) Stath A.) Cross A.) Cross A.) Cross C.) Span A.) Cross A.) Cross A.) Cross C.) Spice A.) Stath A.) Stath	RESULT 7 T13564 microtubule-ass microtubule-ass microtubule-ass microtubule-ass c) Expecies: Dros C; Date: 13-Aug- C; Date: 13-Aug- C; Date: 13-Aug- C; Date: 13-Aug- C; Accession: T1 A; Deference num A; Accession: T1 A; Accessi	RESULT 7 11354 microtubule-associated protein homolog - fruit fly (Drosophila melanogaster NA Alternate names: hypothetical protein EG:49E4.1 C;Species: Drosophila melanogaster C;Date: 13-Aug-1999 #sequence_revision 13-Aug-1999 #text_change 17-Nov-2000 C;Accession: T1356 R;Spanos, L.; Papagianaxis, G.; Siden-Kiamos, I.; Louis, C. Submitted to the EMBL Data Library, April 1999 A;Description: Sequencing the distal X chromosome of Drosophila melanogaste A;Reference number: Z17689 A;Accession: T13564 A;Accession: T13566 A;Accession: T13564 A;Accession: T13564 A;Accession: T13564 A;Accession: T13564 A;Accession: T13564 A;Accession: T13564 A;Accession: T13564 A;Accession: T13564 A;Accession: T13564 A;Accession: T13564 A;Accession: T13564 A;Accession: T13564 A;Accession: T13564 A;Accession: T13564 A;Accession: T13564 A;Accession: T13566 A;Accession: T13564 A;Accession: T13564 A;Accession: T13564 A;A	er) 00 ter.

```
C, Accession: S56231
R; Murakami, Y.; Naitou, M.; Haqiwara, H.; Shibata, T.; Ozawa, M.; Sasanuma, S.I.; Sas submitted to the EMBL Data Library, May 1995
A; Description: Analysis of the nucleotide sequence of chromosome VI from Saccaromyces A; Reference number: S56186
A; Accession: S56231
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        A;Status: preliminary
A;Molecuel Eype: DNA
Residues: 1-796 <MUR>
A;Cross-references: EMBL:D50617; NID:g836685; PIDN:BAA09215.1; PID:d1009856; PID:g836
                                                      34;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            hypothetical protein YFL023w - yeast (Saccharomyces cerevisiae)
C;Species: Saccharomyces cerevisiae
C;Date: 02-Sep-1995 #sequence_revision 19-oct-1995 #text_change 29-oct-1999
                                                                                                                                                                                                                                                                                                                                                                                                 1487 SGIDEKSALASKEASRPESVTDKSKEPSRRESIAESLKAESTKDEKSAPPSKEASRPGSV 1546
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              1759 PASVAESVKDEAEKSKEESRRESVAEKSPLPSKEASRPASVAESIKD-----EAEKS 1810
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 811 KEESRRESV-----AEKSPLPSKEASRPASVAESIKDEAEKSKEE-----SRRESV 1856
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               1857 AEKSPLP-----SKEASRPASVAESIKDEAEKSKEESRRESVAEKSPLPSKEASRPAS 1909
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             2022 -AESIKDEAEKSKEESRRESV------AEKSPLPSKEASRPASVAESIKDEAE 2067
                                                                                                                                                 1701 DGAAQSRETSRPASVAESAKDGADDLKELSRPESTTQSKEAGSIKDEKSPLASE--EASR 1758
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            1910 VAESIKDEAEKSKEESRRESVAEKSPLPSKEASRPASVAESIKDEAEKSKEESRR---- 1964
                                                                                                                                                                                                                                                                                       1427 FASGEASRPESVAESVKDEAGKAESRRES1AKTHKDESSLDKAKEQESRRESLAESIKPE 1486
                                                                                                                                                                                                                                                                                                                                              -----MALEGLARMERRLQRATKRQKKQLEEDGIPVPSVVSRFPTAPYHH----- 132
                                                                                                                                                                                                                                                                                                                                                                                                                                                                 133 - KSTNAEEREAKEP-----VLKTHSKDVE-REAE--IGVDDVVKMEPAATNIIEPEDAQD 183
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            184 A--AERGAARPPAVNSSYLPLPWKGRLGYACLNTYLRNAK-PPIFSSRTCRMASIV---- 236
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      ----DHRHPLQFED--EPEHHLKNKPD----KSKEPQDELGHKFVQEL----- 274
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          -GLANARDIVK--MLCWNEKYGIRFLR------LSSEMFPFASHPVHGYK 315
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       316 LAPFASEVLAEAGRVAAELGHRLTTHPGQFTQLGSPRKEVVESAIRDLEYHDELLSLLKL 375
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       376 PEQQNRDAVMIIHMGGQFGDKAATLERFKRNYARLSQSCKNRLVLENDDVGWTVHDLLPV 435
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        436 CEELNIPMVLDYHHHNICFDPAHLREGTLDISDPKLQERIANTWKRKGIKQKMHYSEPCD 495
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        496 GAVTPRD-----RRKHRPRVMTLPPCPP----DMDLMIEAKDKEQAVFELMRTFKLPG 544
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      545 FEKINDMVPYDRDDENRPAPPV-----KAPKKKKGGKRKRTTD-----EEAAEPEEVDT 593
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         594 AADDVKDAPEGPKEVPEEERAMGGPYNRVYWPLGCEEWLKPKKREVKKGKVPEEVEDEGE 653
                                                                                                                3 SRKSKAAALDTPQSESSTFS----STLDSSAPSPARNLRRSGRNILQPSSE-KDRDHE 55
                                                                                                                                                                                                                                56 KRSGE------ELAGRMMGKD---ANGH----CLREGKEQE-----EGVK-- 87
                                                      Matches 151; Conservative 104; Mismatches 306; Indels 219;
   DB 2; Length 5327;
4.5%; Score 155.5; DE 19.4%; Pred. No. 0.56;
                                 Similarity
      Query Match
                                    Best Local
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        RESULT 8
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         237
                                                                                                                                                                                                                                                                                                                                                    88
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     1602
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                qq
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   q
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              q
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 pp
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  qq
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     QQ
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  qq
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            QY
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             _{\rm Qy}
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             Ω
                                                                                                                                                                                                                                      Qγ
                                                                                                                                                                                                                                                                                          q
                                                                                                                                                                                                                                                                                                                                                    ΟY
                                                                                                                                                                                                                                                                                                                                                                                                                                                                 \Omega
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         Ω
```

```
hypothetical protein DKF2p434F192.1 - human (fragment)
C;Species: Homo sapiens (man)
C;Date: 15-Oct-1999 #sequence_revision 15-Oct-1999 #text_change 15-Oct-1999
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     963 KPARPKEEPEDKAEPAKEKHSPVEERKPIKETAKPAKAAPAKADKEPEA------ 1011
    429 PSAKPKAAKVEKKVVSKKPEIKVESEPISAQLDTDLEDLAQEEVMEAKAAPVVSAEKDEE 488
                                                    --KRSGEELAGRMMGKDANGHCLREGKEQEEGVKMAIEGLARMERRL 100
                                                                                                 489 EEEEEEKEEEEKEEEEAEAEBEEEEEDRGRKEGEAEAEEEAEEEVEKEEAEEAEVEEAEAETEA 548
                                                                                                                                                   -----KQLEEDGI-PVPSVVSRFPTAPYHHKSTNAEEREA 142
                                                                                                                                                                              143 KEPVLKTHSKDVEREAEIGVDDVVKMEPAATNIIEPEDAQDAAERGAARPPAVNSSYLPL 202
                                                                                                                                                                                                                                                                                                                                                                                      -----KAKTQEAEVEEEBABAAEAEEBABEBAGEBDVEABSKEEBEBDSKEA 700
                                                                                                                                                                                                                                                                                                                                                                                                                                        262 -- PQDELGHKFVQELGLANARDIVKMLCWNEKYGIRFLRLSSEMFPFASHPVHGYKLAPF 319
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       ASEVLAEAGRVAAELGHRLTTHPGQFTQLGSPRKEVVESAIRDLEYHDELLSLLKLPEQQ 379
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    722 AEEAEAEAEEEAA----KSEEEAABEAKDEAEEEEABEEAVEETEAATE------EAE 769
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                428 TVHDLLPVCEELNIPM--VLDYHHHNICFDPAHLREGTLDISDPKLOERIANTWKRKGIK 485
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             486 QKMHYSE-----POPCPGAVTPRDRRKHRP---RVMTLPPCPPDMDLMI 524
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             -- RDDENRPAPPVKAPKKK 573
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      915 EKEDDEEEE-------EEVEEVRPEDAKPVKSKPAPAKEEEDEPKPAKQPPKPKR 962
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      574 GGKRKRTTDEEAAEP-----EEVDTAADDVKDAPEGPKEVPEEERAMGGPYNRVYWP 625
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 770 AKEA-----SDDEKPEEEVKESEAPVAPEAKKAPEPKAAPKKAPAKVES---- 814
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             -----PTSEPEDEPKAEVVE-------KKGKAEAPKPKAKPAAAKK-EAKPVE 854
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               Query Match 4.0%; Score 140; DB 2; Length 727; Best Local Similarity 19.4%; Pred. No. 0.38; Matches 131; Conservative 92; Mismatches 213; Indels 238; Gaps
                                                                                                                                                                                                                                                                           203 PWKGRLGYACLNTYLRNAKPPIFSSRTCRMASIVDHRHPLQFEDEPEHHLKNKPDKSKE-
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        C;Accession: T17292
Fabloorge, W.; Wirkner, U.; Mewes, H.W.; Gassenhuber, J.; Wiemann, submitted to the Protein Sequence Database, September 1999
A:Reference number: Z18729
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            A,Cross-references: EMBL:AL117534.
A:Experimental source: adult testis; clone DKFZp434F192 C,Genetics:
A;Note: DKFZp434F192.1
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 380 NRDAVMITHMGGQFGDKAATLERFKRNYARLSQSCKNR----
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        701 DAEEDEAEEEEVKE-----EEVTK------
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                1012 -----AEPKKIEVKVKKVTKKVVEEIE 1033
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    626 LGCEEWLKPKKREVKKGKVPEEVEDEGE 653
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         525 EAKDKEQAVFELMRTFKLPGFEKINDMVPYD-
                                                                                                                                                101 QRATKRQK-----
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         A; Molecule type: mRNA
A; Residues: 1-727 <ANS>
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              Note: DKFZp434F192.1
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         A; Accession: T17292
A; Status: preliminary
                                                 54 HE----
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          RESULT 10
                                                                                                                                                                                                                                                                                            609
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             815
                                                                                                                                                                                                                                                                                                                                                                                         652
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                T17292
    qq
                                               ò
                                                                                            QQ
                                                                                                                                                ò
                                                                                                                                                                                           qq
                                                                                                                                                                                                                                        ŏ
                                                                                                                                                                                                                                                                                       Op
                                                                                                                                                                                                                                                                                                                                         ò
                                                                                                                                                                                                                                                                                                                                                                                         qq
                                                                                                                                                                                                                                                                                                                                                                                                                                        ŏ
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     Оb
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       ŏ
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               Op
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 ΟŊ
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              qq
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           qq
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             ŏ
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          qq
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      ag
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      Qγ
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                q
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  Óγ
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               qq
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      гe
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            R'Jacobs, A.J.; Kamholz, J.; Selzer, M.E.
Brain Res. Mol. Brain Res. 29, 43-52, 1995
A:Title: The single lamprey neurofilament subunit (NF-180) lacks multiphosphorylation
A;Reference number: 151116; MUID:95287814
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           24;
                                                                                                                                                28;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         C;Species: Petromyzon marinus (sea lamprey)
C;Date: 13-Sep-1996 #sequence_revision 13-Sep-1996 #text_change 21-Jul-2000
C;Accession: I51116
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          A,Molecule type: mRNA
A,Residues: 1-1110 <JAC>
A;Cross-references: EMBL:U19361; NID:g632548; PIDN:AAA80106.1; PID:g632549
C;Superfamily: neurofilament triplet H protein
                                                                                                                                                                                                                                        326 KSSFAENSDKKOKKKQVGFASSLEI --- HEVENLKEENKROMOSFAVPMYETQESTG -- I 380
                                                                                                                                                                                                                                                                                       63 AGRMMGKDANGHC-----LREGKEQEEGVKMAIEGLARMERRLQRATKRQKKQLEEDGI 116
                                                                                                                                                                                                                                                                                                                                                                                 117 PVPSVVSRF-----PTAPYHHKSTNAEEREA--KEPVLKTHSKDVEREAEIG--- 161
                                                                                                                                                                                                                                                                                                                                                                                                                  ----VDDVVKMEPAATNIIEPEDA-QDAAERG-------AARPPAVNS-S 198
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               199 YLPLPWKGRLGYACLNTYLRNAKPPIFSSRTCRM---ASIVDHRHPLQFEDEPEHHLKNK 255
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            PDKSKEPQDELGHKF----VQELGLANARDIVKMLCWNEKYGIRFLRLSSEMFPFASHPVH 312
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        313 G-----YKLAPFASEVLAEAGRVAAELGHRLTTHPGQFTQLGSPRKEVVESAIRDLE 364
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     644 GNLDDMVKAYSLGLYDDDLEEDPGTIVEKL----------EDFK 677
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     365 YHDELLSLIKLPEQQNRDAVMIIHMGGQFGDKAATLERFKRNYARLSQSCKNRLVLENDD 424
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       ---EEEN 707
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 425 VGWTVHDLLPVCEELNIPMVLDYHHHNICFDPAHLREGTLDISDPKLQERIANTWKRKGI 484
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             Gaps
                                                                                                                                                                                        4 RKSKAAALDTPQSESST-FSSTLDSSAPSPARNLRRSGRNILQPSSEKDRDHEKRSGEEL 62
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      Indels 239; Gaps
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   2 PSRKSKAAALDTP-----QSESSTFSSTLDSSAPSPARNLRRSGRNILQPSSEKDRD 53
                                                                                                                                                                                                                                                                                                                                    881 ANKMISDEFDGDLFAKMLGVQEADEVHEKYKEELINQERLEGEASRSNRRTR-----
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            545 EFETPFKKK -----KLKSLQKP-RSSKSMKKKFDPKILENISDDDYDDDDGNKKLL
                                                                                                                                         Matches 108; Conservative 97; Mismatches 183; Indels 174;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         Length 1110;
                                                                                              Length 796;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             Best Local Similarity 17.6%; Pred. No. 0.53;
Matches 132; Conservative 102; Mismatches 275;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  678 EYNKQVELLR---DEIRDF------QLENKPVTME-----
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    4.1%; Score 141.5; DB 2; 17.6%; Pred. No. 0.53;
                                                                                            4.2%; Score 146; DB 2; 19.2%; Pred. No. 0.18;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    A; Status: preliminary; translated from GB/EMBL/DDBJ
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             ----SSPAAHTEGELEPIDK 775
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                485 KOKMHYSEPC---DGAVTPRDR 503
                                                                                                                     Best Local Similarity
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                NF-180 - sea lamprey
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              A; Accession: I51116
                      19
                      A; Map position:
                                                                                              Query Match
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       Query Match
C; Genetics
                                                                                                                                                                                                                                                                                                                                                                                                                                   433
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   162
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             260
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     RESULT
                                                                                                                                                                                                                                                                                                                                                                                                                                 g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               g
                                                                                                                                                                                           ò
                                                                                                                                                                                                                                        g
                                                                                                                                                                                                                                                                                       ò
                                                                                                                                                                                                                                                                                                                                  g
                                                                                                                                                                                                                                                                                                                                                                                 ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            음
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            δ
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      q
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     ŏ
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            q
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        ò
```

Οy	62	,	117
οp	21	LENKELGKNGTGG-LFQGKTPLRKANLQQAIVTPLKPVDNTYYKEAEKENLVEQSIPS 7	77
Οy	118		166
QQ	78	NACSSLEVEAAISRKTPAQPQRRSLRLSAQKDLEQKEKHHVKMKAKRCATPVIIDEILPS 1	137
O.y	167	KMEPAATNIIEPEDAODAAERGAARPPAVNSSYLPLPWKGRLGYACLNTYLRNAKP 2 Kili	222 180
3 3	9 6		D 6
g S	181	PITSONICAMASIVORMETAÇE EDEFERINA FUNSA EVO ELGORE V VELGERINAND S P	204
ò	282	GHRLTTH	341
qq	205		222
Οy	342	DELLSLLKLPEQONRDAVMIIHMGGQFGDKAA	398
QQ	223	IGQPVKKSVSQVTKSVDFHFRTDERIKQHPKNQEEYKEV	265
Οy	399	SSCKNRLVLENDDVGWTVHDLLPVCEELNIPMVLDY	447
Dp	266	L: : :: : :: :	314
οy	448	HHHNICFDPAHLREGTLDISDPKLQERIANTWKRKGIKQKMHYSEPCDGAVTPRDRRKHR 5	507
QQ	315		357
Qy	508	PRVMTLPPCPPDMDLMIEAKDKEQAVFELMRTFKLPGFEKINDMVPYD 5	555
ΩD	358	ARAVT CKSTAELEAEELEKLOOYKFKARELDPRILEGGPILPKKPPVKPPTEPIGFD 4	414
οy	556		579
qq	415	LEIEKRIQERESKKKTEDEHFEFHSRPCPTKILEDVVGVPEKKVLPITVPKSPAFALKNR 4	474
Οÿ	580		634
qq	475	IRMPTKEDEEEDEPVVIKAQPVPHYGVPFKP-QIPEARTVEICPFSFDSRDKERQLQK 5	531
Qy	635	KKREVKKGKVPE 646	
qq	532	: : : : : EKKIKELQKGEVPK 545	
RESULT 133100 hypothe hypothe C. Spec C. Acte R. Geist Submit A. Desc: A. Refel A. Stacte A. Sta	RESULT 11 T33100 hypothetic C;Species: C;Cacessio R;Geisel, Submitted A;Reference A;Reference A;Refe	al protein H11E01.3 - Caenorhabditis elegans Caenorhabditis elegans -oct-1999 #sequence_revision 29-oct-1999 #text_change 18- n: T33100 C. ; Wamsley, P. to the EMBL Data Library, May 1998 ion: The sequence of C. elegans cosmid H11E01. n: T33100 n: T33100 pre-liminary; translated from GB/EMBL/DDBJ type: DNA type: DN	Feb-2000 CESP:H11E01.3

DB 2; Length 1510;

4.0%; Score 139.5;

Query Match

```
Gynatery. Screening modes set of the sequence of the sequence of the sequence of the sequence of the sequence of the sequence of the sequence of the sequence of the sequence of the sequence of the sequence of the sequence of the sequence of the sequence of the sequence of the sequence of the sequence of the sequence of the sequence of the sequence of the sequence of the sequence of the sequence of the sequence of the sequence of the sequence of the sequence of the sequence of the sequence of the sequence of the sequence of the sequence of the sequence of the sequence of the sequence of the sequence of the sequence of the sequence of the sequence of the sequence of the sequence of the sequence of the sequence of the sequence of the sequence of the sequence of the sequence of the sequence of the sequence of the sequence of the sequence of the sequence of the sequence of the sequence of the sequence of the sequence of the sequence of the sequence of the sequence of the sequence of the sequence of the sequence of the sequence of the sequence of the sequence of the sequence of the sequence of the sequence of the sequence of the sequence of the sequence of the sequence of the sequence of the sequence of the sequence of the sequence of the sequence of the sequence of the sequence of the sequence of the sequence of the sequence of the sequence of the sequence of the sequence of the sequence of the sequence of the sequence of the sequence of the sequence of the sequence of the sequence of the sequence of the sequence of the sequence of the sequence of the sequence of the sequence of the sequence of the sequence of the sequence of the sequence of the sequence of the sequence of the sequence of the sequence of the sequence of the sequence of the sequence of the sequence of the sequence of the sequence of the sequence of the sequence of the sequence of the sequence of the sequence of the sequence of the sequence of the sequence of the sequence of the sequence of the sequence of the sequence of the sequence of the sequence 
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              IgA-specific metalloendopeptidase (EC 3.4.24.13) NMB0700 [imported] - Neisseria menin
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   1231 DRIDLEAYHNSPTKESTSPVTVNPVEVRVSADEVAEHLASPHQNFEDEPSHLVDESEVPN 1290
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               969 E--RDNRSPTFESSVTMQA---AAPLSPAASDHAEQARE--SPSFERAPSLHSQLSGNLE 1021
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               1022 HDDENSAVVEAGQEPATQSPIPLEQEGRF-ERAASV----NSYQASES-----FENQE 1069
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     1070 ATVVEHHAGALQSPVASEKEVSEAISQPAAEPAAPSPISSEVEAHSEHHHTQEEIPIVQT 1129
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               1130 SEHIDLTDVPRSPAFSMN-----DSHHEDQLSHEAAP----HSPVAFTEPRVP-- 1173
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         1174 ---SVHTSDTEEESFEVQHHAGLTQEYAARKSPTEMDLYNPSRDNQESPVMSETEAIPIV 1230
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            520 MDLMIEAKDKEQAVFELMR----TFKLPGFEKINDMVPYDRDDENRPAPPVKAP----- 569
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        365 YHDELLSELKIPEQQNRDAVMIIHMGGQFGDKAATLERFKRNYARLSQSCKNRLVLENDD 424
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  461 EGTLD1SD-PKLQERIANTWKRKGIKQKMHYSEPCDGAVTPRDRKHRPRVMTLPPCPPD 519
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            ----KDAPE 603
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          821 HSPVASQEEAARSPSVHSSHASEHFEHHEEAQDSPVASQEKAARSPSVHSSHASEDSEHR 880
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             208 LGY-----ACLINTYLRNAKPPIFSSRTCRMASIVDHRHPLQFEDEP-----EHHLKNK 255
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      881 QEIQHSPAASQNEAARS--PSVHSSHASE--HIENHGESLQ---SPVASMAGSEHH--NM 931
                                                                                                                                                                                                                                            | :: | | :: | | 3 | 3 | 4 | 5 | 5 | 5 | 5 | 660 ASERDVPSARSPLVQSENFEHHAEVHSFHASEHIDEALPSPVQSVHSS------ 707
                                                                                                                                                                                                                                                                                                                                                                                                   52 RDHEKRS---GEELAGR---AMGKDANGHCLREGKEQEEGVKMATECLARMERRLQRATK 105
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                708 HDHHDRSSPVASEPAARSPSVQSSRTSEHFEHRGEVPQSPSSNQFHSSEHIEE-----A 761
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   106 RQKKQLEEDG1PVPSVVSRF----PTAPYHHKSTNAEEREAKEPVLKTH-SKDVEREAEIG 161
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          762 ROSPVTNQESVHSPHASEHFEHREVVP-HSPAASQEEFGRSPSVHSPHVSEHFEHHEEAQ 820
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             162 VDDVVKMEPAÅ------TNIIEPEDAQD---AAERGAARPPAVNSSYLPLPWKGR 207
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              256 PDKSKEPQDELGHKFVQELGLANARDIVKMLCWNEKYGIRFLRLSSEMF-----PFAS 308
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           932 AESSEYTTSE------KEISPSI-----FSSHTSEQFEQQSQNSPVAS 968
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        309 HPVHGYKLAPFASEVLAEAGRVAAELGHRLTTHPGQFTQLGSPRKEVVES----AIRDLE 364
; Pred. No. 1.1;
86; Mismatches 302; Indels 237; Gaps
                                                                                                                                                                   --- LDSSAPSPARNLRRSGRNILQPSSEKD 51
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   -----1328
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      604 GPKEVPEEER---AMGGPYNRVYWP---LGCEEWLKPKKREVKKGKVPEEVEDEGE 653
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   1291 NFETVPEQEAYAPSQDAMMTSIYQPGSDTETREW----
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                C.Species: Neisseria meningitidis
A,Variety: strain MC58 serogroup B; strain 81139
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            570 ----KKKKKGGKKKRTTDEEAAEPEEVDTAADDV----
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        425 VGWTVHDL----LPVCEELNIPMVLDY-----
      Best Local Similarity 19.5%;
                                                      Matches 151; Conservative
                                                                                                                                                                   8 AAALDTPQSESSTFSST---
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     qq
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               qq
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                QQ
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   ΟŊ
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             QΫ́
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    \operatorname{QC}
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       QΥ
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    qq
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       Ωy
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     qq
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        ŏ
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               q
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              δŏ
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  δ
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            Q \mathbf{y}
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            οy
```

```
q
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                Db
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   q
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            ΟY
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   Op
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  òγ
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    õ
                                                                                                                                                                                                                                                                                                                                                                                                                ŏ
                                                                                                                                                                                                                                                                                                                                                                                                                                                     q
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      qq
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 \delta
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        qq
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  ΟŅ
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         Op
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     οy
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       qq
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   Qγ
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            q
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      οy
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           Db
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        ó
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             qq
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     Óλ
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            δλ
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               ŏ
A;Cross-references: GB:AE002424; GB:AE002098; NID:g7225923; PIDN:AAF41117.1; PID:g722592
A;Experimental source: serogroup B, strain MC58
KI.Omhol.1, H.: Poulsen, K.: Mogens, K.
Mol. Microbiol. 15, 495-506, 1995
A;Title: Comparative characterization of the iga gene encoding IgAl protease in Neisseri A;Reference number: S61314; MUID:95302961
A;Reference number: S61326
A;Status: preliminary; nucleic acid sequence not shown
A;Molecule type: DNA
A;Molecule type: BNA
A;Residues: 53-548 < LOM>
A;Residues: 53-548 < LOM>
A;Cross-references: EMBL:X82477; NID:g732856; PIDN:CAA57860.1; PID:g732857
                                                                                                                                                                                                                                                                                                                                                                                                                26;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      ELARROGEERKAAELLAKORAEAER-----EAQALAARRKAEAEEAKROAAELAHRO 1240
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      1241 EAE-RKAAELSANQKAAAEAQALAARQQK-----ALARQQEEARKAAELAVKQKAETER 1293
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       1321 ---ARRQEAERQAAELSAKQKAETDREAAESAKRKAEEEEHRQAAQSQPQRRKRRAAPQD 1377
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          1378 YMAASQNRPKRRGRRSTLPAPPSPSFD-----SSAYAAPRALHNPDWYENDYEIP 1428
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  1025 EQVKRQQAEAERKSAELAKQKAEAERERRELATRQKAEQERSSAELARRHEKEREAAELS 1084
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               965 AENRRRVKPAPSPATNTASQAQKATQTDGAQIAKPONIVVAPPSPQANQAEEAKROQAKA 1024
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                KAATL--ERFKRNYARLSQSCKNRLVLENDDVGWTVHDLLPVCEELNIPMVLDYHHNIC 453
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  FDPAHLREGTLDISDPKLQER-----IANTWKRKGIKQ------KMHYSEPCD 495
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     GAVTPRDRRKHRPRVMTLP-PCPPDMDLMIEAKDKEQAVFELMRTFKLPG-FEKINDMVP 553
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        94 ARMERRLQRA-----TKRQKKQLEEDG--IPVPSVVSRFPTAPYHHKSTNAEEREAKE 144
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           145 PVLKTHSKDVEREA-------EIGVDDVVKMEPAATNII-EPEDAQDAAERG 188
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            244 ------FEDEPEHHLKNKPDKSKE---PQDEL------GHKFVQE--LGLANARD 281
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            282 IVKMLCWNEKYGIRFL---RLSSEMFPFASHPVHGYKLAPFASEVLAEAGRVAAELGHRL 338
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               TTHPGQFTQLGSPRKEVVES---AIRDLEYHDELLSLLKLPEQQNRDAVMIIHMGGQFGD 395
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              916 QEGLNLFDASSVQDR-----SRLSVSLANNHVDLGALRYTIKTENGITRLYNPY 964
                                                                                                                                                                                                                                                                                                                                                                                                                  Gaps
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      554 YDRDDENRPAPPV----KAPKKKKGGKRKRTTDEE-----AAEPEEVDTAADDVKDAPEG
                                                                                                                                                                                                                                                                                                                                                                                                                                                         38 RSGRNILQPSSEKDRDHEKRSGEELAGRMMGKDANGH----CLREGKEQEEGVKMAIEGL
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           189 AARPPAVNSSYLPLPWKGRLGYACLNTYLRNAKPPIFSSRTCRMASIVDHRHPLQ----
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  T28657
blackjack protein, microtubule associated - American bird grasshopper
                                                                                                                                                                                                                                                                                                                                                                     Query Match
4.0%; Score 139; DB 2; Length 1815;
Best Local Similarity 18.7%; Pred. No. 1.4;
Matches 125; Conservative 94; Mismatches 268; Indels 180;
                                                                                                                                                                                                                                                              A;Gene: NMB0700
C;Superfamily: IgA-specific metalloendopeptidase
C;Keywords: hydrolase; metalloproteinase
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       1294 KTAELAKORAAAEAAKROOEAROTAEL---
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  1489 NTELHEK 1495
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        605 PKEVPEE 611
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               RESULT 13
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      1189
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               339
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  454
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                396
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    qq
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         Dp
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             셤
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               õ
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  õ
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                         ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        ŏ
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                g
```

```
C;Species: Schistocerca americana (American bird grasshopper)
C;Date: 15-Oct-1999 #sequence_revision 15-Oct-1999 #text_change 11-May-2000
C;Accession: 728657
B;Zachow, K.R.; Bentley, D.
J. Cell Sci. 109, 1497-1507, 1996
A;Title: Blackjack, a novel protein associated with microtubules in embryonic neurons A;Reference number: 220498; MUID:96393048
A;Reference; number: 220498; MUID:96393048
A;Reference; number: 220498; MUID:96393048
A;Reference; number: 220498; MUID:96393048
A;Reference; number: 220498; MUID:96393048
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              30;
                                                                                                                                                                                                                                                                                                                                                         A;Molecule type: mRNA
A;Residues: 1-1547 - CZAC>
A;Cross-references: BMBL:L76606; NID:g1220510; PID:g1220511; PIDN:AAC37266.1
A;Experimental source: embryo; clone library Kai Zinn
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 1268 SPTSPTKASVPSTPRTSTPKRTQVAPPAAKTTAKTTTTTTTTSAKSLSRPTSNV--PSKP 1325
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  1067 VVVFEELQSQL------NYVGIE-----TETEIAEKAEETAEKTFEKKDIS 1106
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      1151 DKEAHKAGDGEDEVLMITKTPPPTPTPASQTELTKDASEMKAESTVSAVAVASAAVAGVA 1210
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             947 AEAVQNQEPSPSAEQTIGDICVPDLHGVPEALRVCVYPMPVEQSSMKDDVELISPSFESQ 1006
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     1007 EEKREALEKDNIHGTDFSTALKLESLSGEQRDMVETTLNEVSNKTELEELNDLTSNADKD 1066
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   1211 ATTTAITSRKEPALKK---QTPLSKDVKKTLVAKDSDKMKPSTKAPVRQSPGTAKAPTKT 1267
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       ----ELGHRLTTHPGQFT-----QLGSPRKEVVESAIRDLEYHDELLSLLKLPEQQNRD 382
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   383 AVMIIHMGGQFGDKAATLERFKRNYAKLSQSCKNRLVLENDDVGWTVHDLLPVCEELNIP 442
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                443 MYLDYHHHNICFDPAHLREGTLDISDPKLQER-IANTWRRKGIKQKMHYSEPCDGAVTPR 501
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      --GLANARDI-----FPFASHP 3.10
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 EELAGRAMGKDANGHCLREGKEQEEGVKMAIEGLARMER------RLQRATKRQKKQLE 112
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             EDGIPVPSVVSRFPTAPYHHKSTNAEE------REAKEPVLKTHSKDVEREAEIGVDDVV 166
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       KDAAEPLEELRKDAAEPLEELRKDAAEPLEELRKDAAEPLEELRKDAAEPLEELRKDAAE 805
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 KMEPAATNIIEP-·EDAQDAAE-----RGAARPPAVNSSYLPLPWKGRLGYACLNTYLR 218
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    806 PLEELRKDAAEPLEELRKDAAEPLEELRKDAAEP--------------LEELRK 845
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          NAKPPIFSSR--TCRMASIVDHRHPLQFEDEPEHHLKNKPDKSKEPQDELGHKFVQEL-- 274
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        DAAEPLEELRKDAAEPAKVVDELH-----GEDHDL-----MEDVNEAGDTKVEKLDT 892
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                893 ITGLTDIRDIETASFQKDTMKL-----NGVESPELNNEMHTVQNITLENRLSFETASPP 946
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  635 RKDAÁBPLEBLRKDAABPLEBLRKDÁABPLEBLRKDAABPLBELRKDAABPLEBLKKDAA 694
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  Gaps
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            4 RKSKAAALDTPQSESSTFSSTLDSSAPSPARNLRRSGRNILQ----PSSEKDRDHEKRSG 59
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     -------EEERAMGGPYNRVYWPLGC
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 AVFELMRTFKLPGFEKINDMVPYDRD------DENRPAPPVKAPKKKGGKRKRTDE
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                4.0%; Score 138; DB 2; Length 1547, 18.2%; Pred. No. 1.4;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              Indels
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              95; Mismatches 313;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            DRRKHRPR-----VMTLPPCPPDMDLMIEAKDKEQ----
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   311 ---VHGYKLAPFASEVLA-------EAGRVAA--
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     584 EAAEPEEVDTAADDVKDAPEGPKEVP
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              Matches 145; Conservative
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     Best Local Similarity
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        Query Match
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           746
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          167
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            846
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               333
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            502
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     09
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             113
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  219
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               695
```

629 EEWLKPKKREVKKGKVP 645

q Ωÿ òγ qu ò qп qq οy ŏδ οy ò A;Cross-references: EMBL:U83913; NID:g3858884; PID:g3858885; PIDN:AAC72432.1 A;Experimental source: strain Balb/C C;Genetics: proliferation potential related protein - mouse C;Species: Mus musculus (house mouse) C;Bate: 11-Jan-2000 #sequence_revision 11-Jan-2000 #text_change 02-Sep-2000 C;Accession: T42727 R;Witte, M.M.; Scott, R.E. submitted to the EMBL Data Library, November 1998 960 KLERTPEKDKIASSTTPAKKIKLNR-----ETGKKIGNAENASTTKEPSEKLESTSSK 1012 1013 IKQEKVKGKAKRVAGSEGSSSTLVDYTSTSSTGGSPVRKSEEKTDTTRTVIKTM---EE 1069 1070 YNNDNTAPAEDVIIMIHVPQSKWDKDDFESEEDVKTTQPIQSVGKPSSII-KNVTTKPS 1128 1129 ATAKYTEKESEQPEKLQKLPKEASHELMQH-ELRSSKGSASSEKGRAKDREHSGSEKDNP 1187 1245 --KVDYDSRDYSSSKRRDERGELARRKDSPPRGKESLSGQKSKLREERDLPKKGAESKKS 1302 1303 NSSP-----PRDKKPHDHKAPYETKRPCEETKPVDKISGKEREKHAAEARNGKDSSGGK 1356 1188 DKRKSGAQPDKESTVDRLSEQGHFKTLSQSSKETRTSEKHESVRGSSNKDFTPGRDK--- 1244 292 YGIRFLRLSSEMPPFASHPVHGYKLAPFASE-----VLAEAGRVAAELGHRLTTHP- 342 61 ELAGRAMGKDANGHCLREGKEQEEGVKMAIEGLARMERRLQRATKRQKKQLEEDGIPVPS 120 798 ATAKK-----KDNVLKPSKGPQEKVDGDREKSPRSEPPLKKA-----KEEATKIDS 842 121 V-----VSRFPTAPYHHKSTNAEEREAK----EPVLKTHSKDVEREAEIGVDDV 165 843 VKPSSSSQKDEKVTGTPRKA--HSKSAKEHQEAKPAKDEKVKKDCSKDIKSEKPASKDEK 900 166 VKMEPAATNII------EPEDAQDAAERGAARPPAVNSSYLPLPWKGRLGYACLN 214 215 TYLRNAKPPIFSSRTCRMASIVDHRHPLQFEDEPEHHLKNKPDKS--KEPQDEL---GHK 269 270 FVQELGLANA-----RDIVKMLCWNEK 291 343 --GQFTQLGSPRKEVVE----SAIRDLEYHDELLSLLKLPEQQNRDAVMIIHMG----- 390 391 -----GQFGDKAATLERFKR--NYARLSQSCKNRLVLENDDV--GWTVHDLLPVCEELNI 441 442 PMVLDYHHHNICFDPAHLREGTL------DISDPKLQERIANTWKRKGIKQKMH 489 749 PSREDATPVRDEPMDAESITFKSVSD------KDKREKDKPKVKSDKTKRKSDGS 797 901 AK-KPEKNKLLDSKGEKRKRKTEEKSVDKDFESSSMKISKVEGTEIVKPSPKRKMEGDVE 959 Gaps 2 PSRKSKAAALDTP-QSESSTFSSTLDSSAPSPARNLRRSGRNILQPSSEKDRDHEKRSGE 60 Matches 157; Conservative 100; Mismatches 326; Indels 232; 3.9%; Score 137; DB 2; Length 1560; 19.3%; Pred. No. 1.6; A; Description: involved in hnRNP association and Rb1 binding C; Superfamily: RING finger homology A;Status: preliminary; translated from GB/EMBL/DDBJ 490 YSEPCDGAVTPRDRRKH------RPRVMTLP--F;57-107/Domain: RING finger homology <RRN> 1326 TSLATQKQSSTTTGKVP 1342 A;Reference number: Z22246 A;Accession: T42727 1-1560 <WIT> Best Local Similarity A; Molecule type: mRNA A; Residues: 1-1560 <W A; Gene: P2P-R Query Match C; Function: δŏ QQ QQ qq δğ g ò QQ δλ g g g QQ Ω δ òγ ò δ g

```
M.; Wohlenberg, C.; Becker, M.H.G.; Key, G.; Flad, H.D.; Ge 2, 1993
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    Rischlueter, C.; Duchrow, M.; Wohlenberg, C.; Becker, M.H.G.; Key, G.; Flad, H.D.; Ge J. Cell Biol. 123, 513-522, 1993
A; Title: The cell proliferation-associated antigen of antibody Ki-67: a very large, u
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  A.Cross-references: EMBL:X6551
C.Suporfamily: kinase interaction domain homology
C.Keywords: alternative splicing, cell cycle control; nucleus; tandem repeat
F:29-91/Domain: kinase interaction domain homology <KIH>
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                C;Species: Homo sapiens (man)
C;Date: 03-May_1994 #sequence_revision 03-May-1994 #text_change 07-May-1999
                                                                           1357 LPCILNPPDLPM-----EKELAVGQVEKSAVKPKPQLSHSSRLSSDLTRETDEAAFEPDY 1411
                                                                                                                                                                                                                                                     1412 NESDSESNVSVKEEERAVASISKDLKEKTTEKAKESLTVATASQPGADRSQSGSPSVSPS 1471
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             2069 KEKAEALEDLVGFKELFQTPGHTEESMTDDKITEVSCKSPQPESFKTSRSSKQRLKIPLV 2128
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   2188 RTRKEKARALEDLVDFKELFS----APGHTEESMTIDKNTKIPCKSPPPELTDTATSTKR 2243
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    2244 CPKTRPRKEVKEELSAVERLTQTSGQSTHTHKEPASGDEGIKVLKQRAKKKPNPVEEEPS 2303
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              2357 TTASTKRHLRTRVQKVQVKEEPSAVKFTQTSG--ETTDADKEPAGEDK-GIKALKESAKQ 2413
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               2414 TPAPAASVTGSRRPRAPRESAQAIEDLAGFKDPAAGHTEESMTDDKTTKIPCKSSPELE 2473
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                2474 DTATSSKRRPRTRAQKVEVKEELLAVGKLTQTSGETTHTDKEPVGEGKGTKAFKQPAKRN 2533
                                                                                                                                                                     559 ENRPAPPVKAPKKKK-----GGKRKRTTDEEAAEPEEVDTAAD----DVKDAPEGPKEVPE 610
515 -PC---PPUMDLM1EAKDKEQAVFELMRTFKLP------GFEKINDMVPYDRDD 558
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  49 EKDRDHEKRSGEELAGRMMGKDANGHCLREG-----KEQEEGVKMAIEGLARME-RRLQR 102
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   147 -----EPEDAQDAAE 186
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             187 RGAARPPAVNSSYLPLPWKGRLGYACLNTYLRNAKPPIFSSRTCRMASIVDHRHPLQFED 246
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         247 --- EPEHHLKNKPDKSKEPQDELGHKFVQELGLANARDIVKMLCWNEKYGIRFLRLSSEM 303
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   304 FPFASHPVHGYKLAPFASEVLAE-----AGRVAAELGH-----RLTTHP----GOFT 346
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    381 RDAVMTIHMGGQ---FGDKAATLERFKRNYARLSQSCKNRLVLENDLJVGWTVHDLLPVCE 437
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 5 KSKAAALD-----TP--QSESSTFSSTLDSSAPSPA----RNLRRSGRNILQPSS 48
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            3.9%; Score 137; DB 2; Length 2897;
Similarity 19.9%; Pred. No. 3.6;
16; Conservative 92; Mismatches 296; Indels 200;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   103 ATKROKKOLEEDGIPVPSVVSRFPTAPYHHKSTNAEEREAKEPV--
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        cell proliferation antigen Ki-67, short form - human
                                                                                                                                                                                                                                                                                                                                                                                                                               1472 RSHSPSGSQTRSHSSSASSAGSQDSKKKKKKKKKEKK 1506
                                                                                                                                                                                                                                                                                                                                     611 EERAMGGPYNRVY----WPLGCEEWLKPKKREVKK 641
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    347 QLGSPRKEVVESAIRDLEYHDELLSLLKL------
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  A; Reference number: A48666; MUID: 94043435
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      =
=
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           A; Molecule type: mRNA
A; Residues: 1-2897 <SCH>
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  A; Status: preliminary
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                C; Accession: B48666
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          A; Accession: B48666
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 146;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                Query Match
Best Local S
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 Matches
```

VDA	ᇤ	2568
	438 ELNIPMVLDYHHHNICFDPAHLREGTLDISDP-KLQERIANTWKRK 482 	482
	PGHTEELANGAADSFTSAPKQTPDSGKPLKISRRVLRAPK 2608	2608
	483 GIKQKMHYSEPCDGAVTPRDRRKHRPRVMT-LPPCPPDMDLMIEAKDKEQAVFELMRTFK 541	541
- 6097	FK	2655
	542 LPGFEKINDMVPYDRDDENRPAPPVKAPKKKGGKRKRTTDEEAAEPEEVDTAADDVKDA 601	601
٠ ٢	2656 MPAPEEIVEELPASKKQRVAPRARGKSSEPVVIMKRSLRTSAKRIEPAE-ELNSNDWKTN 2714	2714
14	602 PEGPKEVPEEE 612	
	2715 KEEHKLQDSVPENK 2728	

Search completed: January 15, 2002, 13:52:31 Job time: 400 sec

	•	
•		
		4.
		•
		. 1
		- 1
		w.'
		•
		J

```
January 15, 2002, 13:49:55; Search time 120.01 Seconds (without alignments) 369.718 Million cell updates/sec
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.
                                                                                                                                                                                                                          US-09-724-296-2_COPY_230_828
3112
1 MLRLLKRNIQISKRIVFTIL......KRLTARKRRSRKEEVEEDEK
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           /SIDS8/goddata/geneseq/geneseqp/AA1981.DAT:*
/SIDS8/goddata/geneseq/geneseqp/AA1982.DAT:*
/SIDS8/goddata/geneseq/geneseqp/AA1982.DAT:*
/SIDS8/goddata/geneseq/geneseqp/AA1984.DAT:*
/SIDS8/gogdata/geneseq/geneseqp/AA1985.DAT:*
/SIDS8/gogdata/geneseq/geneseqp/AA1985.DAT:*
/SIDS8/gogdata/geneseq/geneseqp/AA1987.DAT:*
/SIDS8/gogdata/geneseq/geneseqp/AA1988.DAT:*
/SIDS8/gogdata/geneseq/geneseqp/AA1988.DAT:*
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        /SIDS8/gcgdata/geneseq/geneseqp/AA1980.DAT:*
                                                                                                                                                                                                                                                                                                                                                                                                                                         522463
GenCore version 4.5
Copyright (c) 1993 - 2000 Compugen Ltd.
                                                                                                                                                                                                                                                                                                                                                                                                                                         Total number of hits satisfying chosen parameters:
                                                                                                                                                                                                                                                                                                                                                                                         522463 seqs, 74073290 residues
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              SUMMARIES
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries
                                                                                            OM protein - protein search, using sw model
                                                                                                                                                                                                                                                                                                                         BLOSUM62
Gapop 10.0 , Gapext 0.5
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      Minimum DB seq length: 0
Maximum DB seq length: 200000000
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   A_Geneseq_1101:*
                                                                                                                                                                                                                                    Title:
Perfect score:
                                                                                                                                                                                                                                                                                                                         Scoring table:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     Database :
                                                                                                                                                                                                                                                                                 Sequence:
                                                                                                                                                                                                                                                                                                                                                                                            Searched:
                                                                                                                                            Run on:
```

		фP				
Result No.	Score	Query Match	Query Match Length DB ID	DB	ID	Description
	3112	100.0	828	21	AAY44498	GST signal peptide
10	1979	63.6	371	21	AAY44499	S. pombe delta228-
e	1979	63.6		21	AAY44500	GST signal peptide
4	1011.5	32.5		17	AAW01618	Neurospora crassa
ស	961.5	30.9		21	AAY44502	N. crassa delta228
9	961.5	30.9		21	AAY44503	B. subtilis delta2
7	230.5	7.4		21	AAY44505	D. radiodurans del
80	147.5	4.7		11	AAR06991	Polypeptide antige
σ	140.5	4.5		22	AAB95546	Human protein sequ
10	140.5	4.5	733	19	AAW60785	Deletion mutant de
11	140.5	4.5	2441	21	AAB18161	Plasmodium falcipa

P. falciparum synt Human Acinus L pro Deletion mutant de Deletion mutant de Deletion mutant de Sequence of the Pl Human polypeptide Malarial PERM3 ep Plasmodium falcipa Secreted protein of Human myosin heavy Human polypeptide Saccharomyces cere Amino acid sequenc Human protein sequenc Human protein sequenc Human protein sequenc Human protein sequenc Human protein sequenc Human protein sequenc Human protein sequenc Human protein sequenc Human protein sequenc Peptide #473 enco Peptide #473 enco Peptide #473 enco Peptide #473 enco Peptide #473 enco Peptide #470 enco Arabidopsis thalia	usion protein. endonuclease; UVDE; Uvelp; on damage; fusion protein; tide
AAWS4145 AAWS6128 AAW60787 AAP50777 AAP50777 AAM40833 AAR46605 AAM85608 AAW85608 AAW85608 AAW85608 AAW810529 AAB30809 AAW30296 AAW40137 AAW60795 AAB3348 AAG50809 AAG18695 AAG18695 AAW40137 AAG50808 AAG18695 AAG18695 AAG18695 AAG18695 AAG18695 AAG18695 AAW80795 AAG50808 AAG18695	NMEN agge atio ase bepy aaaa gtt
1639 1341 1753 1654 1658 1658 1668 1753 19 1673 10 10 10 10 10 10 10 10 10 10 10 10 10	T
136.5 134.5 134.5 134.5 134.5 134.5 133.5 134.5 135.5 136.5 137.5 138.5 139.5 13	1498 standa. 1498; AR-2000 (f. signal pept. radiation; cancer; gl. cosaccharom. etic. ide ein difference difference difference difference ide ON 1999;
11111111122222222222222222222222222222	RESULT 1 AAY4449 XX AC AAY4449 XX AC AAY4449 XX CAY4449 X CAY444 XX CAY44 XX CAY444 XX CA

```
ö
                                                                                                                                                                                                                                                                                    The present sequence is a fusion protein comprising the GST signal peptide and full length ultraviolet damage endonuclease (UVDE) from Schizosaccharomyces pombe. UVDE genes are isolated by introducing a foreign cDNA library into a repair-deficient E. coli strain and selecting for complemented cells by UV irradiation of the transformants.
                                                                                                                                                                                                                                                                                                                                                                           This provides stable endonuclease fragments for cleaving a double-stranded DNA molecule that has a distorted structure resulting from UV radiation damage, a photoproduct, an abasic site, mismatched nucleotide pairing, aplatinum diaduct, an intercalated molecule or alkylation of a nucleotide. Uvelp can be used in compositions for internal or topical application and as a therapeutic agent for skin cancers.
                                                                                                                                                                                                new truncated ultraviolet damage endonuclease for treatment of skin
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    TLSMLPQVNIGANSFSAETPVDLKKENETELANISGPHKKSTSTSTRKRARSSKKKATDS 120
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         VSDKIDESVASYDSSTHLRRSSRSKKPVNYNSSSESEEQISKATKKVKQKEEEEYVEE 180
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     VDEKSLKNESSSDEFEPVVPEQLETPISKRRRSRSSAKNLEKESTMNLDDHAPREMFDCL 240
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     410 vdekslknesssdefepvvpeqletpiskrrrsrssaknlekestmnlddhapremfdcl 469
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       350 vsdkidesvasydssthlrrssrskpvnynssseseegiskatkkvkqkeeeyvee 409
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           241 DKPIPWRGRLGYACLNTILRSMKERVFCSRTCRITTIQRDGLESVKQLGTQNVLDLIKLV 300
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             301 EWNHNFGIHFMRVSSDLFPFASHAKYGYTLEFAQSHLEEVGKLANKYNHRLTMHPGQYTQ 360
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             QRLSDSVKARLVLENDDVSWSVQDLLPLCQELNIPLVLDWHHHNIVPGTLREGSLDLMPL 480
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   Gaps
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        361 IASPREVVVDSAIRDLAYHDEILSRMKLNEQLNKDAVLIIHLGGTFEGKKETLDRFRKNY 420
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   grlsdsvkarlvlenddvswsvqdllplcgelniplvldwhhhnivpgtlregsldlmpl 709
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        710 iptiretwtrkgitgkqhysesadptaisgmkrrahsdrvfdfppcdptmdlmleakeke 769
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 1 MLRLLKRNIQISKRIVFTILKQKAFKGNHPCVPSVCTITYSRFHCLPDTLKSLLPMSSKT 60
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          541 QAVFELCRRYELQNPPCPLEIMGPEYDQTRDGYYPPGAEKRLTARKRRSRKEEVEEDEK 599
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          481 IPTIRETWTRKGITQKQHYSESADPTAISGMKRRAHSDRVFDFPPCDPTMDLMIEAKEKE
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       0;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               Length 828;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   Indels
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               DB 21;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         100.0%; Score 3112; DB 21; 100.0%; Pred. No. 1.5e-245;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   .
0
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 Mismatches
                                                                                                                                                                                                                                                   Claim 7; Page 51; 133pp; English.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               ;
                                                                                                        Avery AM;
               98US-0088521.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             Matches 599; Conservative
                                                                                                        Kaur B,
                                                                                                                                           2000-116417/10.
                                                                    (UYEM-) UNIV EMORY.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              Similarity
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     828 AA;
                                                                                                                                                          N-PSDB; AAZ29858
                                  18-MAY-1999;
               08-JUN-1998;
                                                                                                        Doetsch PW,
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         Sequence
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           Query Match
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                Best Local
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        19
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 650
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            121
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    530
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             421
qq
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               Db
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   q
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        á
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   δλ
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         ōλ
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              q
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  a
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        δ
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           δý
```

```
288 amino acid deletion of the N-terminal of the S. pombe uvel+ gene product. This is expressed in frame with a GST leader sequence to generate a fusion protein. This provides stable endonuclease fragments for cleaving a double-stranded DNA molecule that has a distorted structure resulting from UV radiation damage, a photoproduct, an abasic site, mismatched nucleotide pairing, aplathum diadruct, an intercalated molecule or alkylation of a nucleotide. Uvelp can be used in compositions for internal or topical application and as a therapeutic agent for skin
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       The present sequence is delta228-UV damage endonuclease. Delta228-UVDE is
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            A new truncated ultraviolet damage endonuclease for treatment of skin
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               229 DDHAPREMFDCLDKPIPWRGRLGYACLNTILRSMKERVFCSRTCRITTIQRDGLESVKQL 288
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         289 GTQNVLDLIKLVEWNHNFG1HFMRVSSDLFPFASHAKYGYTLEFAQSHLEEVGKLANKYN 348
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   349 HRLTMHPGQYTQIASPREVVVDSAIRDLAYHDEILSRMKLNEQLNKDAVLIIHLGGTFEG 408
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      121 hrltmhpggytgiasprevvvdsairdlayhdeilsrmklneglnkdavliihlggtfeg 180
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             409 KKETLDRFRKNYQRLSDSVKARLVLENDDVSWSVQDLLPLCQELNIFLVLDWHHHNIVPG 468
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 Gaps
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     Delta228-UVDE: ultraviolet damage endonuclease; GST signal peptide; glutathione-S-transferase signal peptide; uvel+ gene product; UV irradiation; DNA damage; UV radiation damage: photoproduct; abasic site; aplatinum diaduct; mismatched nucleotide pairing; nucleotide alkylation; skin cancer.
770 qavfelcrryelqnppcpleimgpeydqtrdgyyppgaekrltarkrrsrkeeveedek
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      63.6%; Score 1979; DB 21; Length 371; 100.0%; Pred. No. 1.9e-153;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 Indels
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    .;
0
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             0; Mismatches
                                                                                                                                                                                                                                                                                                                                                                                   S. pombe delta228-UV damage endonuclease.
                                                                                                                                                                       AAY44499 standard; Protein; 371 AA
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            Claim 13; Page 53; 133pp; English.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        Avery AM;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       99US-0134752.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   99WO-US12910.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         98US-0088521
                                                                                                                                                                                                                                                                                                                (first entry)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          Schizosaccharomyces pombe.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 Conservative
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    Kaur B,
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                2000-116417/10.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              (UYEM-) UNIV EMORY.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               Best Local Similarity
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      371 AA:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     N-PSDB; AAZ29859
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             WO9963828-A1.
                                                                                                                                                                                                                                                                                                                    27-MAR-2000
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       08-JUN-1999;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         08-JUN-1998;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           18-MAY-1999;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                16-DEC-1999.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    Doetsch PW,
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             371;
                                                                                                                                                                                                                                             AAY44499;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      Sequence
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          Query Match
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 cancers
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   cancers.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             Matches
                                                                                                                                     AAY44499
                                                                                                        RESULT
Op
                                                                                                                                                                                                                                             $\circ{\circ}{\circ}$\circ$\circ$\circ$\circ$\circ$\circ$\circ$\circ$\circ$\circ$\circ$\circ$\circ$\circ$\circ$\circ$\circ$\circ$\circ$\circ$\circ$\circ$\circ$\circ$\circ$\circ$\circ$\circ$\circ$\circ$\circ$\circ$\circ$\circ$\circ$\circ$\circ$\circ$\circ$\circ$\circ$\circ$\circ$\circ$\circ$\circ$\circ$\circ$\circ$\circ$\circ$\circ$\circ$\circ$\circ$\circ$\circ$\circ$\circ$\circ$\circ$\circ$\circ$\circ$\circ$\circ$\circ$\circ$\circ$\circ$\circ$\circ$\circ$\circ$\circ$\circ$\circ$\circ$\circ$\circ$\circ$\circ$\circ$\circ$\circ$\circ$\circ$\circ$\circ$\circ$\circ$\circ$\circ$\circ$\circ$\circ$\circ$\circ$\circ$\circ$\circ$\circ$\circ$\circ$\circ$\circ$\circ$\circ$\circ$\circ$\circ$\circ$\circ$\circ$\circ$\circ$\circ$\circ$\circ$\circ$\circ$\circ$\circ$\circ$\circ$\circ$\circ$\circ$\circ$\circ$\circ$\circ$\circ$\circ$\circ$\circ$\circ$\circ$\circ$\circ$\circ$\circ$\circ$\circ$\circ$\circ$\circ$\circ$\circ$\circ$\circ$\circ$\circ$\circ$\circ$\circ$\circ$\circ$\circ$\circ$\circ$\circ$\circ$\circ$\circ$\circ$\circ$\circ$\circ$\circ$\circ$\circ$\circ$\circ$\circ$\circ$\circ$\circ$\circ$\circ$\circ$\circ$\circ$\circ$\circ$\circ$\circ$\circ$\circ$\circ$\circ$\circ$\circ$\circ$\circ$\circ$\circ$\circ$\circ$\circ$\circ$\circ$\circ$\circ$\circ$\circ$\circ$\circ$\circ$\circ$\circ$\circ$\circ$\circ$\circ$\circ$\circ$\circ$\circ$\circ$\circ$\circ$\circ$\circ$\circ$\circ$\circ$\circ$\circ$\circ$\circ$\circ$\circ$\circ$\circ$\circ$\circ$\circ$\circ$\circ$\circ$\circ$\circ$\circ$\circ$\circ$\circ$\circ$\circ$\circ$\circ$\circ$\circ$\circ$\circ$\circ$\circ$\circ$\circ$\circ$\circ$\circ$\circ$\circ$\circ$\circ$\circ$\circ$\circ$\circ$\circ$\circ$\circ$\circ$\circ$\circ$\circ$\circ$\circ$\circ$\circ$\circ$\circ$\circ$\circ$\circ$\circ$\circ$\circ$\circ$\circ$\circ$\circ$\circ$\circ$\circ$\circ$\circ$\circ$\circ$\circ$\circ$\circ$\circ$\circ$\circ$\circ$\circ$\circ$\circ$\circ$\circ$\circ$\circ$\circ$\circ$\circ$\circ$\circ$\circ$\circ$\circ$\circ$\circ$\circ$\circ$\circ$\circ$\circ$\circ$\circ$\circ$\circ$\circ$\circ$\circ$\circ$\ci
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            q
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         ŏ
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   QΥ
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  QQ
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             Qγ
```

```
JP08266275-A.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       29-MAR-1995;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           29-MAR-1995;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        skin cancer.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   15-0CT-1996
                                                                                         Matches 371;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               exposure;
                                                                                                                                                                                                                                                                                                                                                                                                                                                       AAW01618;
molecule
                                        Sequence
                     cancers.
                                                                                                                                                                                              349
                                                                                                                                                                                                                                     409
                                                                                                                                                                                                                                                                              469
                                                                                                                                                                                                                                                                                                                     529
          for
                                                                                                                                                                                                                                                                                                                                                                                                                         AAW01618
                                                                                                                                                                                                                                                                                                                                                                                qq
                                                                                                                                                                                                                                                                                                                                                                                                                                                       qq
                                                                                                                                                                                                                                                                                                 qq
                                                                                                                                                                                                                                                                                                                                          2
                                                                                                                                                                                                                                                                                                                                                              ŏ
22 22 X X QX
                                                                                                               ò
                                                                                                                                 合
                                                                                                                                                      ŏ
                                                                                                                                                                         Οp
                                                                                                                                                                                              δ
                                                                                                                                                                                                                 QQ
                                                                                                                                                                                                                                     ŏ
                                                                                                                                                                                                                                                                              ò
                                                                                                                                                                                                                                                                                                                       δò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      The present sequence is a fusion protein comprising the GST signal peptide and delta228 truncated ultraviolet damage endonuclease (UVDE) from S. pombe. Delta228-UVDE is a 88 amino acid deletion of the N-terminal of the S. pombe uvel+ gene product. This is expressed in frame with a GST leader sequence. Stable endonuclease fragments can be produced for cleaving a double-stranded DNA molecule that has a distorted structure resulting from UV radiation damage, a photopy ...ct, an abasic site, mismatched nucleotide pairing, aplatinum diaduct, an intercalated
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       new truncated ultraviolet damage endonuclease for treatment of skin
                                                                               240
                                                                                                                                                                                                                                                                             ultraviolet damage endonuclease; UVDE; delta228-ÚVDE; repatr-deficient E. coli strain; UV irradiation: DNA damage; UV radiation damage; photoproduct; abasic site; aplatinum diaduct; mismatched nucleotide pairing; nucleotide alkylation;
TMDLMIEAKEKEQAVFELCRRYELQNPPCPLEIMGPEYDQTRDGYYPPGAEKRLTARKRR
                               TLREGSLDLMPLIPTIRETWTRKGITQKQHYSESADPTAISGMKRRAHSDRVFDFPPCDP
                                                                                                                                                                                                                                                                     3ST signal peptide; glutathione-S-transferase signal peptide;
                                                                                                                                                                                                                                                 3ST signal peptide and delta228 S. pombe UVDE fusion protein.
                                                                                                                                                                                                                                                                                                                                                                                           /note= "Encoded by aaaa"
                                                                                                                                                                                                                                                                                                                                                                                                              'note= "Encoded by gtt"
                                                                                                                                                                                                                                                                                                                                                                                                                                  /note= "Encoded by gtt"
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    Disclosure; Page 56; 133pp; English.
                                                                                                                                                                                                                                                                                                                                                                       Location/Qualifiers
                                                                                                                                                                                   AAY44500 standard; Protein; 600 AA.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      Avery AM
                                                                                                                                                                                                                                                                                                                        fusion protein; skin cancer.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                99WO-US12910
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            99US-0134752
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    98US-0088521
                                                                                                                                                                                                                           27-MAR-2000 (first entrý)
                                                                                                                                                                                                                                                                                                                                           Schizosaccharomyces pombe
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      Doetsch PW, Kaur B,
                                                                                                              SRKEEVEEDEK 599
                                                                                                                                 srkeeveedek 371
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         WPI; 2000-116417/10.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 (UYEM-) UNIV EMORY.
                                                                                                                                                                                                                                                                                                                                                                                                      Misc-difference 85
                                                                                                                                                                                                                                                                                                                                                                                 Misc-difference 11
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    N-PSDB; AAZ29860
                                                                                                                                                                                                                                                                                                                                                                                                                         Misc-difference
                                                                                                                                                                                                                                                                                                                                                                                                                                                       WO9963828-A1
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                08-JUN-1999;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             18-MAY-1999;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   08-JUN-1998;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                            16-DEC-1999
                                                                                                                                                                                                                                                                                                                                                      Synthetic
                                                                                                                                                                                                         AAY44500;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  cancers
                                                                                                                                                                 m
                                                                                                                                  361
          181
                                                  241
                                                                                         301
                                                                                                               589
                               469
                                                                      529
                                                                                                                                                                          AAY44500
                                                                                                                                                                 RESULT
                                                                                                                                                                                                          П
                                                                                         g
                                                                                                               ð
                                                                                                                                 염
                               ò
                                                  g
                                                                      ŏ
```

```
.;
0
ecule or alkylation of a nucleotide. Uvelp can be used in compositions internal or topical application and as a therapeutic agent for skin
                                                                                                                                                                                                                                                                                                                                                                                          IMDLMIEAKEKEQAVFELCRRYELQNPPCPLE1MGPEYDQTRDGYYPPGAEKRLTARKRR 588
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           r enzyme; cyclobutane-type dimer; (6-4)-bound product;
UV light; ultraviolet; suntan cream; prevention; treatment;
                                                                                                                                                                                                                                                                   229 DDHAPREMFDCLDKPIPWRGRLGYACLNTILRSMKERVFCSRTCRITTIQRDGLESVKQL 288
                                                                                                                                                                                                                                                                                       289 GTQNVLDLIKLVEWNHNFGIHFMRVSSDLFPFASHAKYGYTLEFAQSHLEEVGKLANKYN 348
                                                                                                                                                                                                                                                                                                                                                                                                                                                           HRLTMHPGQYTQIASPREVVVDSAIRDLAYHDEILSRMKLNEQLNKDAVLIIHLGGTFEG 408
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          KKETLDRFRKNYQRLSDSVKARLVLENDDVSWSVQDLLPLCQELNIPLVLDWHHHNIVPG 468
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                TLREGSLDLMPLIPTIRETWTRKGITOKOHYSESADPTAISGMKRRAHSDRVFDFPPCDP 528
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               530 tmdlmieakekegavfelcrryelgnppcpleimgpeydgtrdgyyppgaekrltarkrr 589
                                                                                                                                                                                                                   Gaps
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               DNA repair enzyme and related DNA - specifically recognises cyclobutane-type dimer and (6-4)-bound product induced by exposure to UV light
                                                                                                                                                                                                                   .;
0
                                                                                                                                                                    DB 21; Length 600;
                                                                                                                                                                                                                   Indels
                                                                                                                                                                                             Pred. No. 3.8e-153;
: Mismatches 0;
                                                                                                                                                                    63.6%; Score 1979;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 Neurospora crassa DNA repair enzyme
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   AAW01618 standard; Protein; 656 AA.
                                                                                                                                                                                                                   0;
                                                                                                                                                                                          100.08;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     95JP-0094137
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    95JP-0094137
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               20-AUG-1997 (first entry)
                                                                                                                                                                                                                      Conservative
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 (SHIS ) SHISEIDO CO LTD.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          SRKEEVEEDEK 599
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            WPI; 1996-512663/51.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                DNA repair enzyme;
                                                                                                                                                                      Query Match
Best Local Similarity
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            Neurospora crassa.
                                                                                                 600 AA;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    N-PSDB; AAT58286
```

```
The present sequence shows a DNA repair enzyme which specifically recognises a cyclobutane-type dimer and (6-4)-bound product induced, respectively from TT and TC sequences in DNA, by exposure to VN light. The enzyme cleaves these sequences at the 5'-side. The DNA repair enzyme can be used in, e.g. suntan cream for prevention and treatment of skin cancer. Conventional endonuclease acts on either the cyclobutane-type dimer or the (6-4)-bound product in damaged DNA, while the enzyme of the invention can act on both of them.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         110 ARSSKKKATDSVSDKIDESVASYDSST-----HLRRSSR-----SKKPVNYNSSSESE- 157
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           ---SEEQISKATKKVKOKEEEEYVEEVDEK 184
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    185 SLKNESSSDEFEPVVPEQLETPISKRRRS-----RSSAKNLEKESTMNLDDHA----- 232
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        115 gipvpsvvsrf-ptapyhhkstnaeereakepvlkthskdvereaeigvddvvkmepaat 173
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     233 ----PREMFDCLDK-------PIPWRGRLGYACLNTILRSMKERVFCSRTCRIT 275
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                --QRDGLESVKQLGTQNVLDLIKLVEWNHNFG 307
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           234 sivdhrhplqfedepehhlknkpdkskepqdelghkfvqelglanardivkmlcwnekyg 293
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                308 IHFMRVSSDLFPFASHAKYGYTL-EFAQSHLEEVGKLANKYNHRLTMHPGQYTQIASPRE 366
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            294 irflrlssemfpfashpvhgyklapfasevlaeagrvaaelghrltthpggftglgsprk 353
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         367 VVVDSAIRDLAYHDEILSRAMKLNEQLNKDAVLIIHLGGTFEGKKETLDRFRKNYQRLSDS 426
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    427 VKARLVLENDDVSWSVQDLLPLCQELNIPLVLDWHHHNIV--PGTLREGSLDLM--PLIP 482
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             483 TIRETWTRKGITQKQHYSESADPTAISGMKRRAHSDRVFDFPPCDPTMDLMIEAKEKEQA 542
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       ----EIMGPE--- 565
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          533 vfelmrtfklpgfekindmvpydrddenrpappvkapkkkkggkrkrttdeeaaepeevd 592
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              3 srkskaaaldtpqsesstfsstldssapsparnlrrsgrnilqpssekdrdhekrsgeel 62
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         -----YYPPGAEKRLTARKRRSRK----EEVEED 597
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               63 agrmmgkdanghclregkegeegvkmaieglarmerrlqratkrgkkgleed-----
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                32.5%; Score 1011.5; DB 17; Length 656; 36.6%; Pred. No. 5.2e-74; Live 88; Mismatches 149; Indels 181;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    N. crassa delta228-UV damage endonuclease.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     543 VFELCRRYEL-------QNPPCP---
        Disclosure; Page 10-13; 18pp; Japanese.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       AAY44502 standard; Protein; 626 AA
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        (first entry)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             Conservative
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           Local Similarity
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      566 --YDOTRDG----
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  656 AA;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           158 -----
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        27-MAR-2000
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     Matches 241;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  Sequence
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                Query Match
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   AAY44502;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            AAY44502
\overset{\mathbf{A}}{\times}\overset{\mathbf{A}}{\times}\overset{\mathbf{A}}{\times}\overset{\mathbf{A}}{\times}\overset{\mathbf{A}}{\times}\overset{\mathbf{A}}{\times}\overset{\mathbf{A}}{\times}\overset{\mathbf{A}}{\times}\overset{\mathbf{A}}{\times}\overset{\mathbf{A}}{\times}\overset{\mathbf{A}}{\times}\overset{\mathbf{A}}{\times}\overset{\mathbf{A}}{\times}\overset{\mathbf{A}}{\times}\overset{\mathbf{A}}{\times}\overset{\mathbf{A}}{\times}\overset{\mathbf{A}}{\times}\overset{\mathbf{A}}{\times}\overset{\mathbf{A}}{\times}\overset{\mathbf{A}}{\times}\overset{\mathbf{A}}{\times}\overset{\mathbf{A}}{\times}\overset{\mathbf{A}}{\times}\overset{\mathbf{A}}{\times}\overset{\mathbf{A}}{\times}\overset{\mathbf{A}}{\times}\overset{\mathbf{A}}{\times}\overset{\mathbf{A}}{\times}\overset{\mathbf{A}}{\times}\overset{\mathbf{A}}{\times}\overset{\mathbf{A}}{\times}\overset{\mathbf{A}}{\times}\overset{\mathbf{A}}{\times}\overset{\mathbf{A}}{\times}\overset{\mathbf{A}}{\times}\overset{\mathbf{A}}{\times}\overset{\mathbf{A}}{\times}\overset{\mathbf{A}}{\times}\overset{\mathbf{A}}{\times}\overset{\mathbf{A}}{\times}\overset{\mathbf{A}}{\times}\overset{\mathbf{A}}{\times}\overset{\mathbf{A}}{\times}\overset{\mathbf{A}}{\times}\overset{\mathbf{A}}{\times}\overset{\mathbf{A}}{\times}\overset{\mathbf{A}}{\times}\overset{\mathbf{A}}{\times}\overset{\mathbf{A}}{\times}\overset{\mathbf{A}}{\times}\overset{\mathbf{A}}{\times}\overset{\mathbf{A}}{\times}\overset{\mathbf{A}}{\times}\overset{\mathbf{A}}{\times}\overset{\mathbf{A}}{\times}\overset{\mathbf{A}}{\times}\overset{\mathbf{A}}{\times}\overset{\mathbf{A}}{\times}\overset{\mathbf{A}}{\times}\overset{\mathbf{A}}{\times}\overset{\mathbf{A}}{\times}\overset{\mathbf{A}}{\times}\overset{\mathbf{A}}{\times}\overset{\mathbf{A}}{\times}\overset{\mathbf{A}}{\times}\overset{\mathbf{A}}{\times}\overset{\mathbf{A}}{\times}\overset{\mathbf{A}}{\times}\overset{\mathbf{A}}{\times}\overset{\mathbf{A}}{\times}\overset{\mathbf{A}}{\times}\overset{\mathbf{A}}{\times}\overset{\mathbf{A}}{\times}\overset{\mathbf{A}}{\times}\overset{\mathbf{A}}{\times}\overset{\mathbf{A}}{\times}\overset{\mathbf{A}}{\times}\overset{\mathbf{A}}{\times}\overset{\mathbf{A}}{\times}\overset{\mathbf{A}}{\times}\overset{\mathbf{A}}{\times}\overset{\mathbf{A}}{\times}\overset{\mathbf{A}}{\times}\overset{\mathbf{A}}{\times}\overset{\mathbf{A}}{\times}\overset{\mathbf{A}}{\times}\overset{\mathbf{A}}{\times}\overset{\mathbf{A}}{\times}\overset{\mathbf{A}}{\times}\overset{\mathbf{A}}{\times}\overset{\mathbf{A}}{\times}\overset{\mathbf{A}}{\times}\overset{\mathbf{A}}{\times}\overset{\mathbf{A}}{\times}\overset{\mathbf{A}}{\times}\overset{\mathbf{A}}{\times}\overset{\mathbf{A}}{\times}\overset{\mathbf{A}}{\times}\overset{\mathbf{A}}{\times}\overset{\mathbf{A}}{\times}\overset{\mathbf{A}}{\times}\overset{\mathbf{A}}{\times}\overset{\mathbf{A}}{\times}\overset{\mathbf{A}}{\times}\overset{\mathbf{A}}{\times}\overset{\mathbf{A}}{\times}\overset{\mathbf{A}}{\times}\overset{\mathbf{A}}{\times}\overset{\mathbf{A}}{\times}\overset{\mathbf{A}}{\times}\overset{\mathbf{A}}{\times}\overset{\mathbf{A}}{\times}\overset{\mathbf{A}}{\times}\overset{\mathbf{A}}{\times}\overset{\mathbf{A}}{\times}\overset{\mathbf{A}}{\times}\overset{\mathbf{A}}{\times}\overset{\mathbf{A}}{\times}\overset{\mathbf{A}}{\times}\overset{\mathbf{A}}{\times}\overset{\mathbf{A}}{\times}\overset{\mathbf{A}}{\times}\overset{\mathbf{A}}{\times}\overset{\mathbf{A}}{\times}\overset{\mathbf{A}}{\times}\overset{\mathbf{A}}{\times}\overset{\mathbf{A}}{\times}\overset{\mathbf{A}}{\times}\overset{\mathbf{A}}{\times}\overset{\mathbf{A}}{\times}\overset{\mathbf{A}}{\times}\overset{\mathbf{A}}{\times}\overset{\mathbf{A}}{\times}\overset{\mathbf{A}}{\times}\overset{\mathbf{A}}{\times}\overset{\mathbf{A}}{\times}\overset{\mathbf{A}}{\times}\overset{\mathbf{A}}{\times}\overset{\mathbf{A}}{\times}\overset{\mathbf{A}}{\times}\overset{\mathbf{A}}{\times}\overset{\mathbf{A}}{\times}\overset{\mathbf{A}}{\times}\overset{\mathbf{A}}{\times}\overset{\mathbf{A}}{\times}\overset{\mathbf{A}}{\times}\overset{\mathbf{A}}{\times}\overset{\mathbf{A}}{\times}\overset{\mathbf{A}}{\times}\overset{\mathbf{A}}{\times}\overset{\mathbf{A}}{\times}\overset{\mathbf{A}}{\times}\overset{\mathbf{A}}{\times}\overset{\mathbf{A}}{\times}\overset{\mathbf{A}}{\times}\overset{\mathbf{A}}{\times}\overset{\mathbf{A}}{\times}\overset{\mathbf{A}}{\times}\overset{\mathbf{A}}{\times}\overset{\mathbf{A}}{\times}\overset{\mathbf{A}}{\times}\overset{\mathbf{A}}{\times}\overset{\mathbf{A}}{\times}\overset{\mathbf{A}}{\times}\overset{\mathbf{A}}{\times}\overset{\mathbf{A}}{\times}\overset{\mathbf{A}}{\times}\overset{\mathbf{A}}{\times}\overset{\mathbf{A}}{\times}\overset{\mathbf{A}}{\times}\overset{\mathbf{A}}{\times}\overset{\mathbf{A}}{\times}\overset{\mathbf{A}}{\times}\overset{\mathbf{A}}{\times}\overset{\mathbf{A}}{\times}\overset{\mathbf{A}}{\times}\overset{\mathbf{A}}{\times}\overset{\mathbf{A}}{\times}\overset{\mathbf{A}}{\times}\overset{\mathbf{A}}{\times}\overset{\mathbf{A}}{\times}\overset{\mathbf{A}}{\times}\overset{\mathbf{A}}{\times}\overset{\mathbf{A}}{\times}\overset{\mathbf{A}}{\times}\overset{\mathbf{A}}{\times}\overset{\mathbf{A}}{\times}\overset{\mathbf{A}}{\times}\overset{\mathbf{A}}{\times}\overset{\mathbf{A}}{\times}\overset{\mathbf{A}}{\times}\overset{\mathbf{A}}{\times}\overset{\mathbf{A}}{\times}\overset{\mathbf{A}}{\times}\overset{\mathbf{A}}{\times}\overset{\mathbf{A}}{\times}\overset{\mathbf{A}}{\times}\overset{\mathbf{A}}{\times}\overset{\mathbf{A}}{\times}\overset{\mathbf{A}}{\times}\overset{\mathbf{A}}{\times}\overset{\mathbf{A}}{\times}\overset{\mathbf{A}}{\times}\overset{\mathbf{A}}{\times}\overset{\mathbf{
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           δ
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               a
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            ζ
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     δλ
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         Op
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              δy
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   qq
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            Op
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         δ
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      qq
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       qq
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                Ω
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            δ
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      δy
```

```
Delta228-UVDE is a 288 amino acid deletion of the N-terminal of the uvel+ gene product. This is expressed in frame with a GST leader sequence to generate a fusion protein. This provides stable endonuclease fragments for cleaving a double-stranded DNA molecule that has a distorted structure resulting from UV radiation damage, a photoproduct, an abasic site, mismatched nucleotide pairing, aplatinum diaduct, an intercalated molecule or alkylation of a nucleotide. Uvelp can be used in compositions for internal or topical application and as a therapeutic agent for skin
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             The present sequence is delta228-UV damage endonuclease from N. crassa.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      A new truncated ultraviolet damage endonuclease for treatment of skin
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     63 agrmmgkdanghclregkegeegvkmaieglarmerrlgratkrgkkgleed----- 114
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          233 ----PREMFDCLDK------PIPWRGRLGYACLNTILRSMKERVFCSRTCRIT 275
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   174 niiepedagdaaergaarppavnssylplpwkgrlgyaclntylrnakppifssrtcrma 233
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 39 evvesairdLeyhdellsllklpeqqnrdavmiihmgqqfqdkaatlerfkrnyarlsgs 398
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        110 ARSSKKKATDSVSDKIDESVASYDSST-----HLRRSSR-----SKKPVNYNSSSESE- 157
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              -----SEEQISKATKKVKQKEEEEYVEEVDEK 184
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             ------QRDGLESVKQLGTQNVLDLIKLVEWNHNFG 307
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  234 sivdhrhplqfedepehhlknkpdkskepqdelghkfvqelglanardivkmlcw---- 288
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            308 IHFMRVSSDLFPFASHAKYGYTL-EFAQSHLEEVGKLANKYNHRLTMHPGQYTQIASPRE 366
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    ------fpfashpvhgyklapfasevlaeagrvaaelghrltthpgqftqlgsprk 338
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              367 VVVDSATRDLAYHDEILSRMKLNEQLNKDAVLIIHLGGTFEGKKETLDRFRKNYQRLSDS 426
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               3 srkskaaaldtpgsesstfsstldssapsparnlrrsgrnildpssekdrdhekrsgeel 62
Delta228-UVDE; ultraviolet damage endonuclease; GST signal peptide; qlutathionerS-transferase signal peptide; uvel+ gene product; UV irradiation; DNA damage; UV radiation damage; photoproduct; abasic site; aplathuum diaduct; mismatched nucleotide pairing; nucleotide alkylation; skin cancer.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                85; Mismatches 145; Indels 181;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           DB 21; Length 626;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          185 SLKNESSSCEFEPVVPEQLETPISKRRS----RSSAKNLEKESTMNLDDHA---
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      30.9%; Score 961.5; DB 2
36.2%; Pred. No. 5.9e-70;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   Claim 16; Page 59; 133pp; English.
                                                                                                                                                                                                                                                                                                                                                                                                         Avery AM;
                                                                                                                                                                                                                                                                                                  98US-0088521.
                                                                                                                                                                                                                                                            99WO-US12910
                                                                                                                                                                                                                                                                                                                          99US-0134752
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  Conservative
                                                                                                                                                                                                                                                                                                                                                                                                         Kaur B,
                                                                                                                                                                                                                                                                                                                                                                                                                                                WPI; 2000-116417/10.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    Query Match
Best Local Similarity
                                                                                                                                                                                                                                                                                                                                                               (UYEM-) UNIV EMORY.
                                                                                                                                 Neurospora crassa.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           626 AA;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           276 TI-----
                                                                                                                                                                         WO9963828-A1.
                                                                                                                                                                                                                                                        08-JUN-1999;
                                                                                                                                                                                                                                                                                                  08-JUN-1998;
                                                                                                                                                                                                                                                                                                                          18-MAY-1999;
                                                                                                                                                                                                                                                                                                                                                                                                         Doetsch PW,
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  Matches 233;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           158 -----
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             Sequence
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      cancers.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              cancers
                        Qγ
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 q
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  q
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               qq
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  q
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            Óζ
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           õ
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          Ö
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  q
```

```
The present sequence is delta228-UV damage endonuclease from B. subtilis. Delta228-UVDE is a 288 amino acid deletion of the N-terminal of the uvel+ gene product. This is expressed in frame with a GST leader sequence to generate a fusion protein. This provides stable endonuclease fragments for cleaving a double-stranded DNA molecule that has a distorted structure resulting from UV radiation damage, a photoproduct, an abasic site, mismatched nucleotide pairing, aplatinum diaduct, an intercalated molecule or alkylation of a nucleotide. Uvelp can be used in compositions for internal or topical application and as a therapeutic agent for skin
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                new truncated ultraviolet damage endonuclease for treatment of skin
427 VKARLVLENDDVSWSVQDLLPLCQELNIPLVLDWHHHNIV--PGTLREGSLDLM--PLIP 482
                  518 vfelmrtfklpgfekindmvpydrddenrpappvkapkkkkggkrkrttdeeaaepeeve 577
                                                                                                                                                                                                                                                                                                                                                                                                                                                                       Delta228-UVDE; ultraviolet damage endonuclease; GST signal peptide; glutathione-S-transferase signal peptide; uvel+ gene product; UV irradiation; DNA damage; UV radiation damage; photoproduct; abasic site; aplatinum diaduct; mismatched nucleotide pairing; nucleotide alkylation; skin cancer.
                                                                                                      459 riantwkrkgikgkmhysepcd-gavtprdrrkhrprvmtlppcppdmdlmieakdkega
                                                                     TIRETWTRKGITQKQHYSESADPTAISGMKRRAHSDRVFDFPPCDPTMDLMIEAKEKEQA
                                                                                                                                                                                                            563 GPEYDQTRDG----YYPPGAEKRLTARKRRSRK----EEVEED 597
                                                                                                                                                                                                                                           578 vpeeeramggpynrvywplgceewlkpkkrevkkgkvpeevede 621
                                                                                                                                                                                                                                                                                                                                                                                                                                        B. subtilis delta228-UV damage endonuclease.
                                                                                                                                          ----ONDPC----
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               Claim 16; Page 59; 133pp; English.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              Avery AM;
                                                                                                                                                                                                                                                                                                                                  AAY44503 standard; Protein;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      98US-0088521
99US-0134752
                                                                                                                                                                                                                                                                                                                                                                                                      (first entry)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              Kaur B,
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             WPI; 2000-116417/10
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         (UYEM-) UNIV EMORY.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               Bacillus subtilis.
                                                                                                                                        VFELCRRYEL-
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 626
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    08-JUN-1999;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                WO9963828-A1
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     08-JUN-1998;
18-MAY-1999;
                                                                                                                                                                                                                                                                                                                                                                                                      27-MAR-2000
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   16-DEC-1999
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              Doetsch PW,
                                                                                                                                                                                                                                                                                                                                                                     AAY44503;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   Sequence
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                cancers
                                                                                                                                                                                                                                                                                                 9
                                                                                                                                          543
                                                                                                                                                                                                                                                                                                                AAY44503
                                                                                                                                                                                                                                                                                                                                                                   ò
                                                                                                      g
                                                                                                                                        ŏ
                                                                                                                                                                       g
                                                                                                                                                                                                            õ
                                                                                                                                                                                                                                             셤
```

Query Match 30.9%; Score 961.5; DB 21; Length 626; Best Local Similarity 36.2%; Pred. No. 5.9e-70; Matches 233; Conservative 85; Mismatches 145; Indels 181; Gaps

```
399 cknrlvlenddvgwtvhd11pvceelnipmvldyhhnnicfdpahlregtldisdpklge 458
                                                                                           63 agrımmgkdanghclregkeqeegvkmaieglarmerrlgratkrqkkqleed----- 114
                                                                                                                             SLKNESSSDEFEPVVPEQLETPISKRRS-----RSSAKNLEKESTMNLDDHA----- 232
                                                                                                                                               233 ----PREMFDCLDK------PIPWRGRLGYACLNTILRSMKERVFCSRTCRIT 275
                                                                                                                                                                                                            308 IHFMRVSSDLFPFASHAKYGYTL-EFAQSHLEEVGKLANKYNHRLTMHPGQYTQIASPRE 366
                                                                                                                                                                                                                                                                                                                                           367 VVVDSAIRDLAYHDEILSRMKLNEQLNKDAVLIIHLGGTFEGKKETLDRFRKNYQRLSDS 426
                                                                                                                                                                                                                                                                                                                                                                                                          VKARLVLENDDVSWSVQDLLPLCQELNIPLVLDWRHHNIV--PGTLREGSLDLM--PLIP 482
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      483 TIRETWTRKGITQKQHYSESADPTAISGMKRRAHSDRVFDFPPCDPTMDLMIEAKEKEQA 542
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     459 riantwkrkgikgkmhysepcd-gavtprdrrkhrprvmtlppcppdmdlmieakdkega 517
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     518 vielmrtiklpgiekindmvpydrddenrpappvkapkkkggkrkrttdeeaaepeeve 577
                                                                                                                                                                                                                                                            -----QRDGLESVKQLGTQNVLDLIKLVEWNHNFG 307
110 ARSSKKKATDSVSDKIDESVASYDSST-----HLRRSSR-----SKKPVNYNSSSESE- 157
                            3 srkskaaaldtpqsesstfsstldssapsparnlrrsgrnilqpssekdrdhekrsgeel 62
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                Delta228-UVDE: ultraviolet damage endonuclease; GST signal peptide; glutathione-S-transferase signal peptide; uvel+ gene product; UV irradiation: DNA damage; UV radiation damage; photoproduct abasic site; aplatinum diaduct; mismatched nucleotide pairing;
                                                                                                                                                                                                                                                                                        234 sivdhrhplgfedepehhlknkpdkskepgdelghkfygelglanardivkmlcw----
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     563 GPEYDQTRDG----YYPPGAEKRLTARKRRSRK----EEVEED 597
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      D. radiodurans delta228-UV damage endonuclease.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                          nucleotide alkylation; skin cancer.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     AA
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     AAY44505 standard; Protein; 294
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                99WO-US12910.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                98US-0088521.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 99US-0134752
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   27-MAR-2000 (first entry)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  Dienococcus radiodurans.
                                                                                                                                                                                                                                                             276 TI----
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 18-MAY-1999;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 WO9963828-A1
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                08-JUN-1998;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 16-DEC-1999
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   AAY44505;
                                                                                                                                                                                                                                                                                                                                                                                                                                                           427
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  AAY44505
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     qq
                                                              Qγ
                                                                                              qq
                                                                                                                              ò
                                                                                                                                                           QQ
                                                                                                                                                                                             QY
                                                                                                                                                                                                                            Op
                                                                                                                                                                                                                                                            ŏ
                                                                                                                                                                                                                                                                                        qq
                                                                                                                                                                                                                                                                                                                            ò
                                                                                                                                                                                                                                                                                                                                                         qq
                                                                                                                                                                                                                                                                                                                                                                                         ò
                                                                                                                                                                                                                                                                                                                                                                                                                         q
                                                                                                                                                                                                                                                                                                                                                                                                                                                           ŏ
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      Db
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      ο̈́O
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      δ
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     qq
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       ò
```

```
q
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       q
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      ΩÝ
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    ŏ
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          δ
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    δ
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             qq
                                                                                                                                                                                                                                                                                                                                                                                           10;
                                                                                                                                                    The present sequence is delta228-UV damage endonuclease from D. radiodurans. Delta228-UVDE is a 288 amino acid deletion of the N-terminal of the uvel+ gene product. This is expressed in frame with a GST leader sequence to generate a fusion protein. This provides stable endonuclease fragments for cleaving a double-stranded DNA molecule that has a distorted structure resulting from UV radiation damage, a photoproduct, an abasic site, mismatched nucleotide pairing, aplatinum diaduct, an intercalated molecule or alkylation of a nucleotide. Uvelp can be used in compositions for internal or topical application and as a therapeutic agent for skin cancers.
                                                                                   A new truncated ultraviolet damage endonuclease for treatment of skin
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   59 rlyrlssslfpmldlagddtgaavlthlapqlleaghaftdagvrllmhpeqfivlnsdr 118
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          218 qedpsvrewvlraratwqppewqvvhls----ngiegpqdrrhshliadfpsayadvpq 272
                                                                                                                                                                                                                                                                                                                                                                                         Gaps
                                                                                                                                                                                                                                                                                                                                                                                                                 249 RLGYACLNTILRSMKERVFCSRTCRITTIQRDGLESVKQLGTQNVLDLIKLVEWNHNFGI 308
                                                                                                                                                                                                                                                                                                                                                                                                                                                                         309 HFMRVSSDLFP---FASHAKYGYTLEFAQSHLEEVGKLANKYNHRLTMHPGQYTQIASPR 365
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 366 EVVVDSAIRDLAYHDEI----LSRMKLNEQLNKDAVLIIHLGGTFEGKKETLDRFRKNY 420
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              119 pevressvramsaharvmdglglartpwn-----llllhggkggrgae----laali 166
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         421 QRLSDSVKARLVLENDDVSWSVQDLLPLCQELNIPLVLDWHHHNIVPGTLREGSLDLMP- 479
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               480 -LIPTIRETWTRKGIT-----QKQHYSESADPTAISGMKRRAHSDRVFDFPPCDPTM-D 531
                                                                                                                                                                                                                                                                                                                                                                                                                                              1 qlglvcltvgpevrfrtvtlsryralspaerea--klldlyssniktlrgaadycaahdi
                                                                                                                                                                                                                                                                                                                                                              DB 21; Length 294;
                                                                                                                                                                                                                                                                                                                                                         7.4%; Score 230.5; DB 21; Length 27.9%; Pred. No. 1.1e-10;
Live 42; Mismatches 140; Indels
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           Polypeptide antigenic to rhostry organelles of Plasmodium.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        Malaria; vaccine; rhoptry organelles.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        AAR06991 standard; protein; 782 AA.
                                                                                                                           Claim 16; Page 60; 133pp; English.
                              Avery AM;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       90EP-0104561
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 (first entry)
                                                                                                                                                                                                                                                                                                                                                                        Best Local Similarity 27.99
Matches 88; Conservative
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       532 LMIEAKEKEQAVFEL 546
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  273 ieveakgkeeaiaal 287
                            Kaur B,
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    Plasmodium falciparum
                                                      WPI; 2000-116417/10
(UYEM-) UNIV EMORY
                                                                                                                                                                                                                                                                                                                  294 AA;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       09-MAR-1990;
                            Doetsch PW,
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                16-JAN-1991
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           26-SEP-1990
                                                                                                                                                                                                                                                                                                                 Sequence
                                                                                                                                                                                                                                                                                                                                                            Query Match
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   AAR06991;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    167
Dp
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   셤
                                                                                                                                                                                                                                                                                                                                                                                                                   ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                           δ
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   ŏ
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         ŏ
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               δ
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       Q
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      Q
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 g
```

```
43 FHCLPDTLKSLLPMSSKTTLSMLPQVNIGANSFSAETPVDLKKENETELANISGPHKKST 102
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       122 fkaspsvvktstpsgtqt-----sglkssspsstkssspsnvksas-phgesn 168
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  103 STSTRKRARSSKKKATDSVSDKIDESV--ASYDSSTHLRRSSRSKKPVNYN----SSSES 156
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         157 ESEEQISKATKK---VKQKEEEEYVEEVD-----EKSLKNESSSDEFEPVVPEQ---LE 204
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               284 nqvkernvagsfsrffsklnpfkkdeviektevskktfsgigfnltdkeakvlgvgatyg 343
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          249 ----RLGYACLNT-----ILRSMKERV--FCSRTCRITTIQRDGLESVKQLGTQNVLDLI 297
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             344 eypetmlyncpnnsnlfdtieslggriidikkresmisttfeggkeclknmg---vldl- 399
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    298 KLVEWNH-----GYTLE----FAQS 335
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     400 ---elndtçckfgtcigsfgehhlrlyefendlfkfhpnidyltladgyklgknhiyels 456
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               336 H-----LEEVGKLAN-----KYNHR-----LTMHPGQ--YTQIASP 364
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              457 hvnfcllnpktleeflkkkeikdlmggddlikykenfdnfmsisitchiesliyddieas 516
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             --VVDSAI----RDLAYHDEILSRMKLN-----EQLN--KDAVLII- 400
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  401 --HLGGTFEGKKETLDRFRKN------YQRLSDSVK-----ARLVLENDDVSW---SV 442
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    575 ryytayalegvcsylehdksqmytelhiynkivdsvryysscfknvivynaiisgiheki 634
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          443 QDLLPLCQELNIPLVLDWHHHNIVPGTLREGSLDLMPLIPTIRETWTRKGITQKQHYSES 502
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       670 hiyfdptvasyayynldrrtmvtlindy------feakkkeltvivsrmktdm--- 716
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 95; Mismatches 241; Indels 219; Gaps
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     -----ekeikpakkysts 669
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    503 ---ADPTAIS----GMKRRAHSDRVFDFPPCDPTMDLMIEAKEKEQAVFELCRRYELQNP 555
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         205 TPISKRRSRSSAKNLEKESTMNLDDHAPREM------FDCLDKPIPWRG----
                                                                                                                                                                                                                                                                                                                                                       Antigenic polypeptide and DNA encoding it - having a determinant cross reactive with those on the rhostry organelles of the merozoite form of the malaria parasite P falciparum.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   expression system using plasmid pMC9, and may then be used for immunisation against malaria. The product may also be used diagnostically to detect Abs directed against the parasite.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          Query Match 4.7%; Score 147.5; DB 11; Length 782; Best Local Similarity 20.8%; Pred. No. 0.0028;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         Gene product may be isolated from a transformed E.coli (Y1088)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       Claim 1; Fig 2a-b; 29pp; English.
                                                                                                        (HOFF ) HOFFMANN-LA ROCHE AG.
89GB-0019064.
89GB-0005857.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        Conservative
                                                                                                                                                                               Scalfe JG;
                                                                                                                                                                                                                                               WPI; 1990-291721/39.
N-PSDB; AAQ06000.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     782 AA;
22-AUG-1989;
14-MAR-1989;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     365 REV----
                                                                                                                                                                               Ridley RG,
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             Matches 146;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         Seguence
```

Indels 143; Gaps

Mismatches 150;

78;

Conservative

Matches 100;

qq ò

```
till-length cDNAs defined in the specification. Where a primer set comprises: (a) an oligo-dT primer and an oligouncleotide comprises set to the complementary strand of a polynucleotide which comprises one of the 5602 nucleotide sequences defined in the specification, where the coligonucleotide comprises at least 15 nucleotides; or (b) a combination of an oligonucleotide comprising a sequence complementary to the complementary strand of a polynucleotide which comprises a 5'-end sequence and an oligonucleotide comprising a sequence complementary to a polynucleotide which comprises a 3'-end sequence. Where the coligonucleotide which comprises a 1-end sequence, where the coligonucleotide comprises a 1-end sequence, where the coligonucleotide comprises a 1-end sequence, where the compliance of the 5'-end sequence, The primer sets can be used in antisense therapy and in gene therapy. The primers are useful for synthesising polynucleotides, particularly full-length cDNAs. The primers are also useful for the proteins encoded by the full-length cDNAs. The primers also useful for the full-length cDNAs. The primers also useful for the full-length cDNAs. The primers also useful for the full-length cDNAs. The primers also useful for the full-length cDNAs. The primers also useful for the full-length cDNAs. The primers also useful for the full-length cDNAs. The primers also useful for the full-length cDNAs. The primers also useful for the full-length cDNAs. The primers also useful for SNAH3628 and cDNAS and SNAH3628 and cDNAS and SNAH3628 and cDNAS and SNAH3628 and cDNAS and SNAH3628 and cDNAS and SNAH3628 and cDNAS and SNAH3628 and cDNAS and SNAH3628 and cDNAS and SNAH3628 and cDNAS and SNAH3628 and cDNAS and SNAH3628 and cDNAS and SNAH3628 and cDNAS and SNAH3628 and cDNAS and SNAH3628 and cDNAS and SNAH3628 and CNAS and SNAH3628
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  AAB95893 represent human amino acid sequences; and AAH13629 to AAH13632 represent oligonucleotides, all of which are used in the exemplification
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   Primer sets for synthesizing polynucleotides, particularly the 5602 full-length cDNAs defined in the specification, and for the detection and/or diagnosis of the abnormality of the proteins encoded by the
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 Human; primer; detection; diagnosis; antisense therapy; gene therapy.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              Yamamoto J;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        present invention describes primer sets for synthesising 5602
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     AAH13633 to AAH18742 represent human cDNA sequences; AAB92446 to
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           Ţ,
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         Saito K,
Otsuki 1
556 PCPLEIMGPEYDQTRDGYYPPGAEKRLTARKRRSRKEEVEE 596
                                                                               717 ---Islgneeskipnd----ksansklatrlmkkfkaeird 750
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    Claim 8; SEQ ID 18167; 2537pp + CD ROM; English.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      Nagai K,
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               nikawa T, Hayashi K,
Wakamatsu A, Naqai K
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       Human protein sequence SEQ ID NO:18167.
                                                                                                                                                                                                                                                                                                                               AAB95546 standard; Protein; 612 AA
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          sogai T, Nishikawa T,
Sugiyama T, Wakamatsı
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        28-JUL-2000; 2000EP-0116126
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       99JP-0248036
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    99JP-0300253.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 11-JAN-2000; 2000JP-0118776.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              2000JP-0183767
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           2000JP-0241899
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             (first entry)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      the present invention.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               (HELI-) HELIX RES INST.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                WPI; 2001-318749/34.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              Isogai T,
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     612 AA;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               full-length cDNAs
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         02-MAY-2000;
09-JUN-2000;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     Homo sapiens
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               EP1074617-A2
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       29-JUL-1999;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               27-AUG-1999;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         07-FEB-2001
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             26-JUN-2001
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          Sequence
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           Ishii S,
                                                                                                                                                                                                                                                                                                                                                                                                                              AAB95546
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              ï
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        The
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   ota
                                                                                                                                                                                                                                                                                       AAB95546
                                                                                                                                                                                                                                          RESULT
```

Score 140.5; DB 22; Length 612; Pred. No. 0.0072;

4.5%;

Query Match Best Local Similarity

```
153 rrvaq1qrqied1kgdeakaket1kkyegeirqleealvhvrk--eekeavsarralene 210
                                                                                                                                             RRRSRSSAKNLEKESTM----NLDDHAPREMFDCLDKPIPWRGRLGYACLNTILRSMKER 265
                                                                                                                                                                                                                      -----laemqrqlkek 306
                                                                                                                                                                                                                                                   VFCSRTCRITTIQ-RDGLESVK-----QLGTQNVLDLIKLVEWNHNFGIHFMRVSSDLF 318
                                                                                                                                                                                                                                                                                                                  ----HRLTMHPG 356
                                                                                                                                                                                                                                                                                                                                                                               357 QYTQIASPREVVVDSAIRDLAYHDEILSRMKLNEQLNKD-AVLIIHLGGTFEGKKETL-- 413
48 DTLKSLLPMSSKTTLSMLPQVNIGANSFSAETPVDLKKENETELANISGPHKKSTSTSTR 107
                             eatknvevlasrsntsegdga----gtemrvkllgeenekl-----grseele 152
                                                            108 KRARSSKKKATDSVSD--KIDESVASYDSS-----THIRRSSRSKKPVNYNSSSESE 157
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  Elongin A; rat; elongation activity; stimulation; RNA polymerase II; binding; Elongin B; Elongin C; regulation; transcription activity; RNA polymerase; in vitro transcriptional assay; deletion mutant.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       Isolated EP and C and fragments - used for regulating transcriptional activation of RNA polymerase, particularly for use
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     AAW60783-92 represent deletion mutants created to define Elongin A
                                                                                                                           SEE---QISKATKKVKQ-----KEEEEYVEEVDEKSLKNESSSDEFEPVVPEQLETPISK
                                                                                                                                                                                                                                                                                                                                                                                                                                                                460 gmeariaeledrleseerdranlglsnrrlerkvk-elvmgyddehlsltd 509
                                                                                                                                                                                                                                                                                                                                                                                                                                             ----DRFRKNYQ----RLSDSVKARLVLENDDVSWSVQD
                                                                                                                                                                                                                                                                                                                    319 PFASHAKYGYTLEFAQSHLEE----VGKLANKYN----
                                                                                                                                                                                                                       263 lqkemadiveasrtstlelqnqldeykeknrre----
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     Deletion mutant delta481-520 of Elongin A.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      Disclosure; Pages 80-83; 185pp; English
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            (OKLA-) OKLAHOMA MEDICAL RES FOUND.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          AAW60785 standard; protein; 733 AA.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          in in vitro assays and systems
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              97WO-US17992
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             96US-0725459
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         (first entry)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             Conaway
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           WPI; 1998-286420/25
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       08-SEP-1998
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 W09814574-A1
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               02-OCT-1997;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             04-OCT-1996;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             Conaway JW,
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               09-APR-1998
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    Synthetic.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     Rattus sp.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          AAW60785;
                                                                                                                                                                                                                                                                                                                                                                                                                                               414 ---
                                108
                                                                                                                           158
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          qq
                                                                                                                                                                                                                                                                                                                                                                                                                qq
                               qq
                                                                                            qq
                                                                                                                                                         q
                                                                                                                                                                                                                      Db
                                                                                                                                                                                                                                                        δ
                                                                                                                                                                                                                                                                                    gg
                                                                                                                                                                                                                                                                                                                                                                                                                                               δy
   ó
                                                                ò
                                                                                                                           ò
                                                                                                                                                                                          õ
```

```
15;
sequences required for binding to Elongins B and C and for transcriptional activity. Elongin A stimulates the elongation activity of RNA polymerase II. Mutants that span the transcriptional activation domain from residues 400 to 730 (AAAW60775) were constructed and expressed in Escherichia coli. The present deletion mutant lacks amino acids 481 to 520 of the wild type Elongin A (AAR60771). The Elongin polypeptides and fragments can be used for regulating the transcriptional activity of RNA polymerase. They can be used in in vitro transcriptional
                                                                                                                                                                                                                                                                                                                                                                                                            ekd---regnslkkklspaldvasdnhfkkpkhkdsekiksdknkqsvdsvd----s 346
                                                                                                                                                                                                                                                                                                                                                                                                                                                                          Gaps
                                                                                                                                                                                                                                                                                                        73 NSFSAETPVDLKKENETELANISGPHKKSTSTSTRKRARSSKKKATDSVSDKIDESVASY 132
                                                                                                                                                                                                                                                                                                                                          243 sshkekrpvdargdekssvmgrekshkasskeesrrll----sedsakeklpssvvkk 296
                                                                                                                                                                                                                                                                                                                                                                            133 DSSTHLRRSSRSKKPVN--YNSSSESEEQISKATKKVKQKEEEEYVEEVDEKSLKNES 190
                                                                                                                                                                                                                                                                                                                                                                                                                                                  191 SSDEFEPV-----VPEQLETPISKRRS--RSSAKNLEKESTMNLDDHAPR-----EMF 237
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       238 DCLDKPIPWRGRLGYACLNTILRSMKERVFCSRTCRITTIQRDGLESVKQLGT-QNVLDL 295
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           -----rkkkkkvvktsg---talgekglkkkdskstsknlnsa 447
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           297 IKLVEWNHNFGIHFM------RVSSDLFPFA------SHAKYGYTLEF 332
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              448 qklpkanenksdklqpagaeptrprkvptdvlpeagftgrrmnskmqvysgskcaylpkm 507
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  333 AQSHLEEVGKLANKYNHRLTMHPGQYTQIASPREVVVDSAIRDLAYHDEILSRMKLNEQL 392
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  508 mtlhqqcirvlknnidsifevggvpysvl----epvlerctpdqlyrieecnhvliee-- 561
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         Plasmodium falciparum; chromosome 2; human malaria parasite; vaccine; antimalarial; malaria; protozoacide; infection; insecticide.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     Plasmodium falciparum chromosome 2 related protein SEQ ID NO:18.
                                                                                                                                                                                                                                                                        83;
                                                                                                                                                                                                                                       DB 19; Length 733;
                                                                                                                                                                                                                                                                     75; Mismatches 167; Indels
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      393 NKDAVLIIHLGGTF-EGKKETLDRFRKNYQRLSDSVKARLVLENDDV 438
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            4.5%; Score 140.5; DB 20.1%; Pred. No. 0.0094
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              AAB18161 standard; Protein; 2441 AA.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       99WO-US26796
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           98US-0107131
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    (first entry)
                                                                                                                                                                                                                                                                       82; Conservative
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              Plasmodium falciparum
                                                                                                                                                                                                                                                    Best Local Similarity
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            (GARD/) GARDNER M. (VENT/) VENTER J C.
                                                                                                                                                                                  733 AA;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           HOFFMAN S.
CARUCCI D.
                                                                                                                                              assays or systems
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           407 lsydqp-----
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                WO200025728-A2.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       05-NOV-1999;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         05-NOV-1998;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    07-NOV-2000
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   11-MAY-2000.
                                                                                                                                                                                  Sedneuce
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    AAB18161;
                                                                                                                                                                                                                                     Query Match
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           (HOFF/)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               (CARU/)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      ٩
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            RESULT 11
                                                                                                                                                                                                                                                                       Matches
                                                                                                                                                                                                                                                                                                                                                                                                              297
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               AAB18161
     δ
                                                                                                                                                                                                                                                                                                                                          q
                                                                                                                                                                                                                                                                                                                                                                          ò
                                                                                                                                                                                                                                                                                                                                                                                                            q
                                                                                                                                                                                                                                                                                                                                                                                                                                                  δ
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       ð
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         a
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             δ
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             Ω
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                ŏλ
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  pp
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      ŏ
```

```
The present invention describes proteins and their fragments (I) encoded by chromosome 2 of the human malarial parasite, Plasmodium falciparum. Also described are: (1) nucleotide sequences (II) encoding (I); and (2) vaccines against P. falciparum infection comprising (I) or (II).

(I) and (II) are useful for the development of vaccines against. P. falciparum infection. (I) and polyclonal antisera or a monoclonal antibody raised to immunogens comprising the sequences of (I), are useful in the detection of infection with P. falciparum. Furthermore, (I) (especially when they are rifins or secreted or membrane proteins) or an aid the identification of drugs to treat or prevent P. falciparum of infection, or they can be used to identify drug resistance in P. falciparum. Sequencing of the Plasmodium chromosome 2 and the complexity of the parasite biology, a process hampered by the complexity of the parasite biology, a process hampered by the complexity of the parasite lifecycle, and provide new targets for ancience of a drug accine and drug development. Parasite resistance to drugs and mosquito
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       parts of the world, and there is a pressing need for vaccines and new drugs. AAA70078 to AAA70078 and AAA18144 to AAB18352 represent nucleotide and protein sequences given in the present invention, but which are not specifically mentioned within the specification.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          resistance to insecticides have led to a resurgence of malaria in many
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     162 I-SKATKKVKOKEEEETVEEVDEKSLKNESSSDEFEPVVPEQLETPISKRRRSRSSAKNL 220
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            55 PMSSKTTLSMLPQVNIGANSFSA-----ETPVDLKK-----ENETELANISGPHKKST 102
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  55 eirlkrslagv----lgntrlssrgvrdprtkealkekgfrdhkrkealkgkteknekar 110
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           111 nalkekklkegkkndagkakdltkke--sqdss----sekslkekvngealkekenket 163
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               221 EKESTMNLDDHAPREMFDCLDKPIPWRGRLGYACLNTILRSMKERVFCSRTCRITTIQRD 280
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         224 ikg---hlkdyeerk----ekrrnw-----ilrslr-----rd 249
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        281 GLESVKQLGTQNVLDLIKLVEWNHNFGTHFMRVSSDLFPFASHAKYGYTLEFAQSHLEEV 340
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 ----eklnaqlesainelke- 273
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 341 GKLANKYNHRLTMHPGQYTQIASPREVVVDSAIRDLAYHDEILSRMKLNEQLNKDAVLII 400
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 274 -rrasrrpmmvkmgrgmkde-----vdewik--kyddegaekngtkdeeikdk---- 318
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                401 HLGGTFEGKKET------LDRFRKNYQRLSDSVKARLVLENDDVSWSVQDLLPL 448
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          --gdgyeeivetkfygmrenalgeldeyeerye-----kkryylk-edgegdlkd---- 365
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            103 STSTRKRARSSKKKATDSVSDKIDESVASYDSSTHLRRSSRS-KKPVNYNSSSESEEEQ 161
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         449 CQELNIPLVLDWHHHNIVPGTLREGSLDLMPLIPTIRETWTRKGITQKQHYSESADPTAI 508
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              1 MLRLLKRNIQISKRIVFTILKQKAFKGNHPCVPSVCTITYSRFHCLP-----DTLKSLL 54
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               | : | : | : | | : : | | | : : | | : : | | : : | | : : | | : : | | : : | | : : | | : : | | : : | | : : | | : : | | : : | | : : | | : : | | : : | | : : | | : : | | : : | | : : | | : : | | : : | | : : | | : : | | : : | | : : | | : : | | : : | | : : | | : : | | : : | | : : | | : : | | : : | | : : | | : : | | : : | | : : | | : : | | : : | | : : | | : : | | : : | | : : | | : : | | : : | | : : | | : : | | : : | | : : | | : : | | : : | | : : | | : : | | : : | | : : | | : : | | : : | | : : | | : : | | : : | | : : | | : : | | : : | | : : | | : : | | : : | | : : | | : : | | : : | | : : | | : : | | : : | | : : | | : : | | : : | | : : | | : : | | : : | | : : | | : : | | : : | | : : | | : : | | : : | | : : | | : : | | : : | | : : | | : : | | : : | | : : | | : : | | : : | | : : | | : : | | : : | | : : | | : : | | : : | | : : | | : : | | : : | | : : | | : : | | : : | | : : | | : : | | : : | | : : | | : : | | : : | | : : | | : : | | : : | | : : | | : : | | : : | | : : | | : : | | : : | | : : | | : : | | : : | | : : | | : : | | : : | | : : | | : : | | : : | | : : | | : : | | : : | | : : | | : : | | : : | | : : | | : : | | : : | | : : | | : : | | : : | | : : | | : : | | : : | | : : | | : : | | : : | | : : | | : : | | : : | | : : | | : : | | : : | | : : | | : : | | : : | | : : | | : : | | : : | | : : | | : : | | : : | | : : | | : : | | : : | | : : | | : : | | : : | | : : | | : : | | : : | | : : | | : : | | : : | | : : | | : : | | : : | | : : | | : : | | : : | | : : | | : : | | : : | | : : | | : : | | : : | | : : | | : : | | : : | | : : | | : : | | : : | | : : | | : : | | : : | | : : | | : : | | : : | | : : | | : : | | : : | | : : | | : : | | : : | | : : | | : : | | : : | | : : | | : : | | : : | | : : | | : : | | : : | | : : | | : : | | : : | | : : | | : : | | : : | | : : | | : : : | | : : | | : : | | : : | | : : | | : : | | : : | | : : | | : : | | : : | | : : | | : : | | : : | | : : | | : : | | : : | | : : | | : : | | : : | | : : | | : : | | : : | | : : | | : : | | : : | | : : | | : : | | : : | | : : | | : : | | : : | | : : | | : 
                                                                                                                                                                                                                                      Proteins encoded by chromosome 2 of the human malarial parasite, Plasmodium falciparum, useful as antimalarial vaccines and in the
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          DB 21; Length 2441;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     Query Match 4.5%; Score 140.5; DB 21; Length Best Local Similarity 19.0%; Pred. No. 0.056; Matches 125; Conservative 104; Mismatches 218; Indels
                                                       Venter JC;
                                                                                                                                                                                                                                                                                                                                                                                                                          Disclosure; Page 50-57; 577pp; English.
                                                                                                                                                                                                                                                                                                                                     diagnosis of P. falciparum infection -
                                                       Gardner M,
                                             Carucci D,
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  250 klreieql-----
                                                                                                                                            WPI; 2000-365347/31.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   2441 AA;
                                             Hoffman S,
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   Sequence
q
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            ŏ
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      qq
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                qq
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         q
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          ŏλ
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               Ωý
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             ŏ
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             qq
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            q
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           δ
```

op O	366	
ογ δ	509 SGMKRRAHSDRVFDFPPCDPTMDLMIEAKEKGQAVFELCRRYEL 552	
a o	QNPPCPLEIMGPEYDQTRDGYYPPGAEKRLTARKRRSRKEEVEEDE 59	
염 .	457 selfpivprkdnelavsgdcmdskvngkklkstfnpfkrrrnklkerkmqe 507	
RESULT AAW54145 ID AAW	JT 12 1145 AAW54145 standard; Protein; 1639 AA.	
X S X	AAW54145;	
X L X	23-SEP-1998 (first entry)	
X DE	P. falciparum synthetic gp190 protein.	
KW	<pre>gp190: malaria; MSP-1; merozoite surface protein; stability; vaccine; monoclonal antibody; passive immunisation; parasite.</pre>	
xx os os	Plasmodium falciparum. Synthetic.	
X Z	WO9814583-A2.	
X C	09-APR-1998.	
P.F.	02-0CT-1997; 97WO-EP05441.	
PR X	02-ocT-1996; 96DE-4040817.	
X A X	(BUJA/) BUJARD H.	
Y I X	Bujard H, Pan W, Tolle R;	
DR DR	WPI; 1998-240088/21. N-PSDB; AAV21451, AAV35363.	
X L L L	Recombinant production of complete gp190/MSP-1 Plasmodium surface protein - useful in anti-malaria vaccines, also stabilising genes by reducing their AT content	
XX PS	Example 1; Fig 3c; 48pp; German.	
\$ 8888888	This sequence represents a modified Plasmodium falciparum gp190/MSP-1 (merozoite surface) protein. The gene encoding this protein has been stabilised by reducing the AT content of the nucleotide sequence. Such a protein is useful in vaccines against malaria or for producing monoclonal antibodies (for passive immunisation). The complete gp190 protein can now be produced outside the parasite and has, at least over extended regions, the capture of follows the parasite and has.	
3883	produced recombinantly than would be possible using the parasites as source.	
SO	Sequence 1639 AA;	
D O N	Query Match 4.4%; Score 136.5; DB 19; Length 1639; Best Local Similarity 19.5%; Pred. No. 0.066; Matches 130; Conservative 107; Mismatches 212; Indels 217; Gaps 35;	
Oy Dp	17 FTILKQKAFKGNHPCVPSVCTITYSRFHCLPDTLKSLLPMSSKTTLSMLPQVNIGAN 73 : :	
λο	SFSAETPVDLKKENETELANISGPHKKSTSTSTRKRARSSKKKATDSVSDKIDESVASYD	
QQ		

```
Chromatin aggregation induction; human; Acinus; AIDS; herpes; diabetes; Alzheimer's disease; Parkinson's disease; cardiac infarction; brain infarction; adenovirus infection; viral hepatitis.
                                                                                                                                                                                                                                                                                                                                                                            701 sepstegeitggattkpgggagsalegd---svgagagagkga-----gppvpv 746
                                                                                                                                                                                                                        ECKKETLDRFRKNYQ--RLSDSVKAR----LVLENDDVSWSVQDLLPLCQELNIPLVLDW 460
                                                                                                                                                                                                                                                                                                                        ADPTA---ISCMKRRAHSDRVFDFPPCDPTMDLMIEAKEKEQAVFELCRRYELQNPPCPL 559
              SLKN------PISKRRRSRSA 217
                                                                    ://
324 eiknpppansgntpntlldknkkieehekeikeiaktikfnidslftdpleleyylrekn 383
                                                                                                 218 KNLE-----KESTMNLDDHAPREMFDCLDKPIPWRGRLGYACLNTILRSM----- 262
                                                                                                                        384 knidisakvetkest-----epneypngvtyp-----isyndinnalnelnsfgdlin 431
                                                                                                                                                    263 ------GTQNVLDLIKLVE-- 301
                                                                                                                                                                  302 ------WNHNFGIHFMRVSSDLFPFAS--HAKYCYTLEFAQSHLEEVGKLANKYNHRL 351
                                                                                                                                                                                                                                                         352 TMHPGQYTQIASPREVVVDSAIRDLAYHDEILSRMK-----LNEQLNKDAVLIIHLGGTF 406
                                                                                                                                                                                                                                                                                                                                                             461 HHHNI-VPGTLREGS------LD----LD----LMPLIPTIRETWTRKGITQKQHYSES 502
134 SSTHLRRSSRSKKPVN-----YNSSSESEEGISKATKKVKQKEEEETVEEVDEK-- 184
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        AAY85657 standard; Protein; 1341 AA.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    Tsujimoto Y;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    Human Acinus L protein sequence.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        07-APR-2000; 2000WO-JP02254
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 99JP-0103317
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           12-FEB-2001 (first entry)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           (SHIO ) SHIONOGI & CO LTD.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    Eguchi Y,
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            WPI; 2000-665130/64.
N-PSDB; AAC61196.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                   560 EIMGPE 565
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 : 11
747 pv--pe 750
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       WO200061743-A1
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               Homo sapiens.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 09-APR-1999;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               19-0CT-2000.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    Sahara S,
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 AAY85657;
                                                  185
                                                                                                                                                                             432
                                                                                                                                                                                                                                                                                                          407
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       qq
                                                                                                                                                                                                                                                                              qq
                                                                                                                                                                                                                                                                                                                                  q
                                                                                                                                                                                                                                                                                                                                                                                    qq
                                                                                                                                                                                                                                                                                                                                                                                                                                        QQ
                      q
                                                 ò
                                                                        qq
                                                                                                 Q
                                                                                                                         QQ
                                                                                                                                                    QΥ
                                                                                                                                                                            qq
                                                                                                                                                                                                     Οy
                                                                                                                                                                                                                                                         οy
                                                                                                                                                                                                                                                                                                         Qy
                                                                                                                                                                                                                                                                                                                                                              δ
                                                                                                                                                                                                                                                                                                                                                                                                                ŏ
                                                                                                                                                                                                                                                                                                                                                                                                                                                                   \overset{\circ}{\circ}
```

```
A novel polypeptide which induces chromatin aggregation but is not involved in DNA fragmentation used for the regulation of apoptosis and screening of potential apoptosis inhibitors.

Disclosure; Page 60-67; 99pp; Japanese.

This invention relates to a polypeptide which induces chromatin aggregation. The polypeptide is a fragment of the human Acinus protein. Included in the invention is a nucleotide sequence encoding the polypeptide, antisense oligonucleotides, antisense or antisense or pulypeptide, nucleotide sequence, and sense or antisense polypeptide, nucleotide sequence, and sense or antisense oligonucleotides. Also included in the invention is a method for screening compounds for their activity as chromatin aggregation inhibitors, regulators or promoters. The polypeptide induces chromatin aggregation within the cell nucleus without inducing DNA fragmentation. The polypeptide can be used in the investigation and treatment of diseases in which apoptosis is implicated, including AIDS, Alzheimer's adenovirus infection, diabetes and viral hepatitis. The present sequence represents the human Acinus protein.
```

Sequence 1341 AA;

```
966 hevkkvtlgdtltrrsisgqksgvsitiddpvrtaqvpspprgkisnivhisnlvrpftl 1025
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         1026 gqlkellgrtgtlveeafwidkikshcfvtystveeavatrtalhg--vkwpqsnpkflc 1083
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    1084 adyaeqdeldyhrgllvd----- 1111
                                                                                                                                                                                                                                                                                                                                                                                                       640 rsrsrssassnsrkslspgv-srdsstsytetkdpssggevatppvpglqvcepkerts 698
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         404 ------GTFEGKKETLDRFRKN----YQRLSDSVKARLVLENDDVSWSVQDLLPLC 449
                                                                                                                                                                       529 lplkieelalakgiteeclkqpsleqkegrrashtllps-----hrlkqsadssss 579
                                                                                                                                                                                                                                   56 MSSKTTLSMLPQVNIGANSFS--AETPVDLKKENETEL---ANISG-PHKKSTSTS-TRK 108
                                                                                                                                                                                                                                                                                            580 rssssssssssrsrspdssgsrshsplrskqrdvagarthanprgrpkmgsrstsesrs 639
                                                                                                                                                                                                                                                                                                                                                                                                                                                                         143 RSKKPVNYNSSSESESEEQISKATKKVKQKEEEEYVEEVDEKSLKNESSSDEFEP----VV 199
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     699 tssssvqarrlsqpesae----khvtqrlq-----pergspkkceaeeaeppaatq 745
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              200 PEQLETPISKRRSRSSAKNLEKESTMNLDDHAPREMFDCLDKPIPWRGRLG------ 251
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       746 pqtsetqtshlpeserihhtveekeevtmdtsenrpendvpeppmpiadqvsnddrpegs 805
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            806 vedeekkesslpksfkrkisvvsatkgvpagnsdtegggpgrkrrwgastattgkkpsis 865
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           283 ---ESVKQL------GTQNVLDL----IKLVE-------WNHNFGIHFMRVSSDLF 318
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 866 ittes1ks1ipdikplaggeavvdlhaddsrisedeterngddgthdkglkicrtvtgvv 925
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             319 PFASHAKYGYTLEFAQSHLEEVGKLANKYNHRLTMHPGQYTQIASPREVVVDSAIRDLAY 378
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         926 paegg-----enggreeeeeek-----epeaeppv--ppgvsvevalpppae 965
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        ---IIHLG----- 403
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           450 QELNIPLVLDWHHHNIVPGTLREGSLDLMPLIPTIRETWTRKGITQKQHYSESADPTAIS 509
                                                                                                                 2 LRLLKRNIQISKRIVFTILKQKAF-----KGNHPCVPSVCTITYSRFHCLPDTLKSLLP 55
                      Best Local Similarity 18.9%; Pred. No. 0.065;
Matches 142; Conservative 107; Mismatches 266; Indels 238;
DB 21; Length 1341;
                                                                                                                                                                                                                                                                                                                                                    109 RARSSKKKATDSVSDKIDESVASYDSSTHL------
4.3%; Score 135;
18.9%; Pred. No. 0
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 379 HD-----EILSRMKLNEQLNKDAVL-----
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      ----YACLNTILRSMKERVFCSRTCR-----
Query Match
                                                                                                                 ò
                                                                                                                                                                       g
                                                                                                                                                                                                                                   δλ
                                                                                                                                                                                                                                                                                         qq
                                                                                                                                                                                                                                                                                                                                                    ŏ
                                                                                                                                                                                                                                                                                                                                                                                                             q
                                                                                                                                                                                                                                                                                                                                                                                                                                                                         ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           δy
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    d
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   δ
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               В
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     Qγ
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             Dp
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             δă
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               a
```

```
16;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         AAW60783-92 represent deletion mutants created to define the Blongin A sequences required for binding to Blongins B and C and for transcriptional activity. Blongin A stimulates the elongation activity of RNA polymerase II. Mutants that span the transcription activation domain from residues 400 to 730 (AAW60775) were constructed and expressed in Escherichia coli. The present deletion mutant lacks amino acids 566-585 of the wild type Elongin A (AAW60771). The Blongin polypeptides and fragments can be used for regulating the transcriptional activity of RNA polymerase. They can be used in in vitro transcriptional assays or
                                     1112 giprplhpppp---ppvqppghpraegregeravregwaereremerrertrserewdrd 1168
510 GMKRRAHSDRVFDFPPCDPTMDLMIBAKEKEQAVFELC--RRYELQNPPCPLEIMGPEYD 567
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         73 NSFSAETPVDLKKENETELANISGPHKKSTSTSTRKRARSSKKKATDSVSDKIDESVASY 132
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 243 sshkekrpvdargdekssvmgrekshkasskeesrrll----sedsakeklpssvvkk 296
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           133 DSSTHLRRSSRSKKPVN--YNSSSESEEQISKATKKVKOKEEEETVVEEVDEKSLKNES 190
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  ekd---regns1kkk1spaldvasdnhfkkpkhkdsekiksdknkgsvdsvd----s 346
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            191 SSDEFEPV-----VPEQLETPISKRRRS--RSSAKNLEKESTMNLDDHAPR-----EMF 237
                                                                                                                                                                                                                                                                                                                                                                                           Elongin A; rat; elongation activity; stimulation; RNA polymerase II; binding; Elongin B; Elongin C; regulation; transcription activity; RNA polymerase; in vitro transcriptional assay; deletion mutant.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              Isolated Elongin A and C and fragments - used for regulating transcriptional activation of RNA polymerase, particularly for use
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          Length 753;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                75; Mismatches 173; Indels
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          DB 19;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        4.3%; Score 134.5; DI 19.8%; Pred. No. 0.03;
                                                                                                                      1169 kvreg -- - prsrsrsrdrrrkerakskekksek 1198
                                                                              568 QTRDGYYPPGAEKRLTARKRRSR-KEEVEEDEK 599
                                                                                                                                                                                                                                                                                                                                                      Deletion mutant delta566-585 of Elongin A.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 Disclosure; Pages 88-91; 185pp; English.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          (OKLA-) OKLAHOMA MEDICAL RES FOUND.
                                                                                                                                                                                                                           AAW60788 standard; protein; 753
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              n in vitro assays and systems
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    96US-0725459.
                                                                                                                                                                                                                                                                                                          08-SEP-1998 (first entry)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  83; Conservative
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   Conaway RC;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           WPI; 1998-286420/25.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              Best Local Similarity
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           753 AA;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          W09814574-A1.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           02-OCT-1997;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  04-OCT-1996;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   09-APR-1998
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   Conaway JW,
                                                                                                                                                                                                                                                                                                                                                                                                                                                                              Synthetic.
                                                                                                                                                                                                                                                                     AAW60788;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               Sequence
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          Query Match
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        systems.
                                                                                                                                                                                      RESULT 14
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                Matches
                                                                                                                                                                                                       AAW60788
                                                                                                                                                                                                                                                                     Öλ
```

ð g ŏ

ò qq δ g

18;

95; Gaps

Mismatches 169; Indels

73;

Conservative

Matches

```
AAW60783-92 represent deletion mutants created to define the Elongin A sequences required for binding to Elongins B and C and for transcriptional activity. Blongin A stimulates the elongation activity of RNA polymerase II. Mutants that span the transcriptional activation domain from residues 400 to 730 (AAW60775) were constructed and expressed in Escherichia coli. The present deletion mutant lacks amino acids 546 to 565 of the wild type Elongin A (AAW60771). The Elongin polypeptides and fragments can be used for regulating the transcriptional activity of RNA polymerase. They can be used in in vitro transcriptional
                                                                                                              238 DCLDKPIPWRGRLGYACLNTILRSMKERVFCSRTCRITTIQRDGLESVKQLGT-QNVLDL 296
                                                                                                                                                                                                                                                                     297 IKLVEWNHNFGIHFM------RVSSDLFPF-----ASHAKYG--YTLEFAQS-- 335
                                                                                                                                                                                                                                                                                                                        ------HLEEVGKLANKYNHRLTMHPGQYTQIASPREVVVDSAIRDLAYHDEILSR 385
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          MKLNEQLN-----KDAVLIIHLGGTF-EGKKETLDRFRKNYQRLSDSVKARLVLENDDV 438
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 568 lyrieecnhvlieetdqlwkvhchrdfkeerpeeyeswremylrlqdaregrlrlltnni 627
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                Elongín A; rat; elongation activity; stimulation; RNA polymerase II;
binding; Elongín B; Elongín C; regulation; transcription activity;
RNA polymerase; in vitro transcriptional assay; deletion mutant.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    Isolated Elongin A and C and fragments - used for regulating transcriptional activation of RNA polymerase, particularly for use in in vitro assays and systems
:|: | | :| | |: | | 347 grgtgdplprakdkvpnnlkaqegkvrtnsdrkspgslpkveemdmddefeqptmsfesy
                                                                                                                                                                                      -----rkkkkvvktsg---talgekglkkkdskstsknlnsa
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   508 pkrkafsspqeeeeagftgrrmnskmqvysgskcaylpkmmtlhqqcirvlknnidsidq
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         Deletion mutant delta546-565 of Elongin A.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           Disclosure; Pages 86-88; 185pp; English.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           (OKLA-) OKLAHOMA MEDICAL RES FOUND
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   A
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           AAW60787 standard; protein; 753
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        97WO-US17992
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               96US-0725459
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  (first entry)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          Conaway JW, Conaway RC;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             WPI; 1998-286420/25.
                                                                                                                                                                                          407 lsydqp-----
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        02-OCT-1997;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               04-OCT-1996;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      WO9814574-A1
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  08-SEP-1998
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  Rattus sp.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           Synthetic.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      AAW60787;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        RESULT 15
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            AAW60787
```

```
73 NSFSAETPVDLKKENETELANISGPHKKSTSTSTRKRARSSKKKATDSVSDKIDESVASY 132
                        :| : ||| : : : : : || | : : : : || sshkekrpvdargdekssvmgrekshkasskeesrfll-----sedsakeklpssvvkk 296
                                                                    133 DSSTHLRRSSRSKKPVN--YNSSSESESEQISKATKKVKQKEEEEYVEEVDEKSLKNES 190
                                                                                                                                           SSDEFEPV-----VPEQLETPISKRRS--RSSAKNLEKESTMNLDDHAPR-----EMF 237
                                                                                                                                                                             347 grgtgdplprakdkvpnnlkagegkvrtnsdrkspgslpkveemdmddefegptmsfesy 406
                                                                                                                                                                                                                DCLDKPIPWRGRLGYACLNTILRSMKERVFCSRTCRITTIQRDGLESVKQLGT-QNVLDL 296
                                                                                                                                                                                                                                                -----rkkkkvvktsg---talgekglkkkdskstsknlnsa 447
                                                                                                                                                                                                                                                                                       297 IKLVEWNHNFGIHFM------RVSSDLFPF-----ASHAKYG--YTLEFAQS-- 335
                                                                                                                                                                                                                                                                                                          pkrkafsspqeeeeagftgrrmnskmqvysgskcaylpfevggvpysvlepvlerctpdq 567
                                                                                                                                                                                                                                                                                                                                                                                                                               377 AYHDEILSRMKENEQLNKDAVLTTHLGGTF-EGKKETLDRFRKNYORLSDSVKARLVLEN 435
                                                                                                                                                                                                                                                                                                                                                                                                                                                  Search completed: January 15, 2002, 13:49:59
                                                                                                                                                                                                                                                  407 lsydqp---
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        DDV 438
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         627
                                                                                                                                             191
                                                                                                                                                                                                                  238
                                                                                                                                                                                                                                                                                                                                                                                               508
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        436
                                                                                                                                                                                                                                                                                                                                                              336
                                                                                                                                                                                                                                                                                                                                                                                             qq
                                                                                                                                                                                                                                                                                                                                                                                                                                   ŏ
   ò
                                 QQ
                                                                        ò
                                                                                                        qq
                                                                                                                                           Óγ
                                                                                                                                                                               QQ
                                                                                                                                                                                                                  \overset{\sim}{\circ}
                                                                                                                                                                                                                                              Dp
                                                                                                                                                                                                                                                                                       ^{\circ}
                                                                                                                                                                                                                                                                                                                         qq
                                                                                                                                                                                                                                                                                                                                                              Óλ
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        ŏλ
```

Job time: 288 sec

Score 134.5; DB 19; Length 753; Pred. No. 0.03;

4.3%;

Query Match Best Local Similarity

753 AA;

Sequence

assays or systems

.

e t